THE

INTERNATIONAL SERIES

OF

MONOGRAPHS ON PHYSICS

GENERAL EDITORS

R H FOWLER AND P. KAPITZA

THE INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS

GENERAL EDITORS

R H FOWLER
F R S , Fellow of Trinity College, Cambridge, Professor of Applied Mathematics in the University of Cambridge

P KAPITZA
F R S , Fellow of Trinity College, Cambridge, Royal Society Messel Research Professor

Already Published

CONSTITUTION OF ATOMIC NUCLEI AND RADIOACTIVITY By G GAMOW 1931 Royal 8vo, pp 122

'This is an important book which will be of great technical use to workers in the subject and in addition provide an informative and interesting account for more general readers'—*Cambridge Review*

THE THEORY OF ELECTRIC AND MAGNETIC SUSCEPTIBILITIES By J H VAN VLECK 1932 Royal 8vo, pp 396

'A most valuable contribution to the study of magnetic and electric susceptibilities and to quantum mechanical theory'—*Nature*

WAVE MECHANICS ELEMENTARY THEORY By J FRENKEL 1932 Royal 8vo, pp 286

'A lucid and elegant summary of the main results of wave mechanics'—*Times Literary Supplement*

WAVE MECHANICS ADVANCED GENERAL THEORY By J FRENKEL 1934 Royal 8vo, pp 533

'Professor Frenkel's work has been eagerly awaited by specialists in physics, and justifies the keenest anticipations by its massive completeness —*Times Literary Supplement*

THE THEORY OF ATOMIC COLLISIONS By N F MOTT and H S W MASSEY 1933 Royal 8vo, pp 300

'There is, we think, no collision problem of any importance which is not mentioned in the book, and for most of them at least a qualitative theory is given   The book will, without doubt, be indispensable for everybody doing research on collisions either theoretical or experimental —*Nature*

RELATIVITY, THERMODYNAMICS, AND COSMOLOGY By R C TOLMAN 1934 Royal 8vo, pp 518

'The book can be most confidently recommended to anyone who wishes to follow the latest developments of this remote, but enthralling, branch of science'—*Cambridge Review*

ELECTROLYTES By HANS FALKENHAGEN Translated by R P BELL 1934 Royal 8vo, pp 364

CHEMICAL KINETICS AND CHAIN REACTIONS By N SEMENOFF 1935 Royal 8vo, pp. 492

THE PRINCIPLES OF QUANTUM MECHANICS By P A M DIRAC Second Edition 1935 Royal 8vo, pp 312

RELATIVITY
GRAVITATION
AND
WORLD-STRUCTURE

OXFORD UNIVERSITY PRESS
AMEN HOUSE, E C 4
LONDON EDINBURGH GLASGOW
NEW YORK TORONTO MELBOURNE
CAPETOWN BOMBAY CALCUTTA
MADRAS SHANGHAI
HUMPHREY MILFORD
PUBLISHER TO THE UNIVERSITY

PLATE I

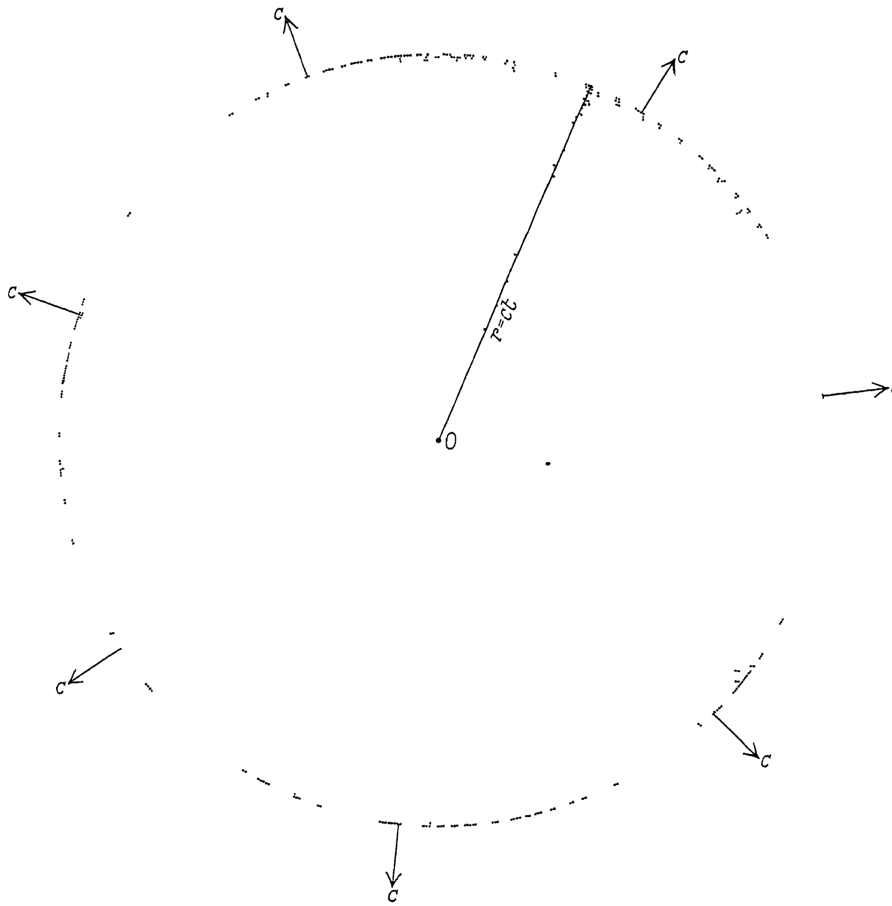


Diagram representing the expanding universe of nebulae, as made by the observer at O, at a particular epoch in his experience Each dot represents the nucleus of a nebula, and is in outward motion from O with uniform velocity At any one epoch, in the experience of O, the velocities are proportional to the distances from O The points are scattered with increasing density farther and farther away from O, the density approaching infinity near the boundary The boundary, which is receding from O in every direction with the speed of light, is not itself occupied by points, but the points form an 'open' set of which every point of the boundary is a limiting point The assembly of moving points has the property that an observer on any point also sees himself as the geometrical centre of the system in his view, with the other points distributed round him with spherical symmetry inside a radius defined by the epoch in his experience to which the diagram refers For each particle-observer, the system is locally homogeneous near himself, the density increasing outwards, at first slowly, ultimately to infinity The total population of points is infinite The particles near the boundary tend towards invisibility, as seen by the central observer, and fade into a continuous background of finite intensity

RELATIVITY GRAVITATION AND WORLD-STRUCTURE

BY

E. A. MILNE
M A, D Sc, F R S
ROUSE BALL PROFESSOR OF MATHEMATICS AND FELLOW OF WADHAM COLLEGE, OXFORD
PAST FELLOW OF TRINITY COLLEGE, CAMBRIDGE

OXFORD
AT THE CLARENDON PRESS
1935

'In the beginning God created the heaven and the earth'

GEN I 1

'Except ye become as little children, ye shall not enter into the kingdom of heaven' MATT XVIII 3

PRINTED IN GREAT BRITAIN

PREFACE

THE present volume contains the material originally delivered as lectures at the University College of Wales, Aberystwyth, on the Aberystwyth Lectures Foundation, in 1933, under the general title 'World-structure problems of time and space and the distribution of matter and motion in the universe' The titles of the five lectures were 'The cosmological problem', 'The earlier gravitational solutions of Einstein and others', 'The kinematic solution', 'The primeval ground-plan of the universe', 'The final picture creation, evolution, gravitation, light' It was part of the conditions of the lectureship that the lectures should be published I desire to express my appreciation of the abundant hospitality I received on the occasions of their delivery from the Principal and members of the staff of the University College of Wales, Aberystwyth

Since the original delivery the material has been considerably enlarged An immature account of the investigations had been previously published in *Zeitschrift fur Astrophysik*, Band 6, S 1–95, 1933,† and I desire to thank the editors of this journal for their courtesy in printing so long a contribution, as well as Professor E F Freundlich for his sympathetic welcome of the ideas After the lectures it became necessary to develop the whole treatment in various directions, so that the present version differs both in spirit and in detail from the earlier account, it also goes much farther A summary of the revised and extended treatment was delivered as a lecture before the London Mathematical Society on May 17th, 1934, under the title 'World-gravitation by kinematic methods', and a more extended treatment was given at the St Andrews' Mathematical Colloquium, held under the auspices of the Edinburgh Mathematical Society, July 1934 An account of some aspects of the subject was given as the Joule lecture for the Manchester Literary and Philosophical Society on February 27th, 1934, under the title 'The expanding universe as a thermodynamic system', and published in the *Memoirs and Proceedings* of that Society, vol 78, March 1934, and a brief treatment of the subject-matter of Chapters II and VII was contained in an address delivered before the British Institute of Philosophy, October 17, 1933, and published in *Philosophy*, January 1934

† Correction slip, Heft 3 of the same volume For preliminary account, *v* *Nature*, 2 July 1932

In the section dealing with Newtonian cosmology I desire to acknowledge the collaboration of Dr W H. McCrea, who is, however, in no way responsible for the line of thought I adopt I desire also to acknowledge the considerable contribution made by Dr A G Walker in the part dealing with the principle of least action, and also certain assistance from Mr. G J Whitrow of Christ Church, Oxford, in one of the developments of Chapter II

As the work is intimately concerned with the limitations of the so-called 'general' theory of relativity, I desire to express my profound admiration for the magnificent work of Professor Einstein. My treatment is essentially different from that developed by him in his Rhodes lectures at Oxford in 1931, but I had the great privilege, on the occasion of his then visit to Oxford, of hearing an exposition by him of some of his ideas in personal conversation But more than that, the whole possibility of the present treatment derives from Einstein's earlier epoch-making work in the subject called 'special' relativity It is impossible to state adequately the inspiration afforded by Einstein's pioneer methods

I wish on this occasion to acknowledge my deep indebtedness to others, in particular to my teachers Mr C H. Gore, late head master of Hymers, and Mr Clifford Chaffer, to Dr E. W. Barnes, my director of studies at Trinity College, Cambridge, to Professor G. H Hardy for teaching me the elements of mathematical rigour and more especially for interesting me in the theory of sets of points (here used) in lectures at Cambridge, to Mr E Cunningham, for his lectures on relativity, to Professor S Chapman, who introduced me to the power of vector analysis, to Professor A. V Hill, who taught me the elements of scientific research, and who convinced me that mathematics, whilst a good servant, is a bad master; to Professor H F Newall, who gave me my first opportunities of studying astrophysics, and has ever since given unfailing encouragement, and lastly to Lord Rutherford, for his insistence on the necessity for obtaining physical insight into physical phenomena If this volume contains anything of value, it is due to the teaching of these friends

WADHAM COLLEGE, OXFORD, *July*, 1934 E A M.

To Professor R H Fowler, my collaborator and friendly critic on so many occasions, I owe useful criticism whilst this book was passing through the press I am also indebted to Mr R G Collingwood for certain suggestions.

December, 1934 E A M

CONTENTS

This volume incorporates the substance of lectures delivered at the University College of Wales, Aberystwyth, on the Aberystwyth Lectures Foundation, March 6th–10th, 1933

I

SCOPE OF THE INVESTIGATIONS

1. The investigations here to be described originated in an attempt to gain insight into the phenomenon of the expanding universe. Through researches begun and continued by V. M. Slipher at the Lowell Observatory, and further researches undertaken by E. Hubble at the Mount Wilson Observatory, it is now well known that the extra-galactic nebulae show spectra in which the lines are displaced to the red, and for good reasons, which we discuss later, this phenomenon is commonly interpreted as indicating that the nebulae are receding from us. It was found by Hubble that the velocities of recession calculated from the observed shifts interpreted as Doppler effects are proportional to the distances of the corresponding objects. The nebulae are therefore separating from ourselves with velocities proportional to their separations from us, and hence are similarly separating from one another. This phenomenon is known as the expansion of the universe. It implies that the mean density of the smoothed-out universe of nebulae is diminishing in any neighbourhood as time goes on. The existence of the phenomenon of the expansion raises questions as to its origin, its rate, and its probable future, and as to its relation to the mean density of the universe, the evolution of individual nebulae, and the evolution of the universe as a whole.

These questions have usually been discussed by means of the concept of 'expanding space'. The concept of 'curved space', previously the province of geometry alone, entered physics with Einstein's general theory of relativity, which gave also an account of gravitation. As gravitational phenomena certainly occur, and as they could be described in terms of the properties of curved space, it came to be believed that 'space' in the neighbourhood of a distribution of matter was 'really curved'. It was a natural step for Einstein to take when he assigned a space of uniform curvature for the purpose of the description of the world as a whole. Einstein convinced himself that a solution of the problems raised by the structure of the universe of matter was impossible within the domain of Newtonian mechanics. But he evaded their difficulties only by the addition of a quasi-empirical term to his law of gravitation, over and above his replace-

ment of the Euclidian space of Newtonian theory by a spherical space. Such a spherical space was modified by Friedmann, in a pioneer paper, by making its curvature change with the time. Lemaître, independently, developed the same idea and showed that his results were in harmony with the observed recession-phenomenon.

Since these pioneer investigations, with which important papers by de Sitter should be included, a host of writers have explored the different possible models that can be constructed for the universe on the basis of general relativity. It has been shown that models can be described in spaces of positive, negative, or zero curvature, expanding, oscillating, or contracting, and with positive, negative, or zero values of the cosmical constant. No criterion has been suggested which would decide which of these would be expected to apply to the universe disclosed by astronomy, nor has any decision been come to as to which of these (if any) does in fact correspond to the observed universe.

It was with no feeling of hostility to these theories, still less with any hostility to the general theory of relativity, that I began to think of the problem for myself. As mathematical investigations the existing theories are of the highest value. But their current explanations in terms of 'expanding space' left me entirely puzzled. Motion imposed in consequence of a geometry differing from the geometry commonly used in physics was a credible notion. Gravitation as a warping of space was a credible notion, though it gave not the least hint as to the nature or origin of gravitation, why the presence of matter should affect 'space' was left entirely unexplained. In assigning structure to space, in restoring structure to structurelessness, mathematical physicists had in effect reintroduced an ether.† They had attributed gravitation to the properties of this ether in relation to the matter present, but had thrown little insight into the situation.

Also Einstein's law of gravitation, amazingly successful and satisfying as it was, was by no means an inevitable consequence of the conceptual basis given by describing the phenomena by means of a Riemannian metric. I had never been convinced of its necessity, and when one reflected that there was no more virtue in the value $\lambda = 0$ of the 'cosmical constant' λ than in any other value, its essential arbitrariness became more apparent. Again, what was it that was

† This is illustrated by Sir James Jeans's remark that the great nebulae are 'mere straws floating in the stream of space', which 'ought to show us in what way, if any, its currents are flowing' (*The New Background of Science*, 1933, p. 137).

curved, when it was stated that space is curved? Still more, what was it that was 'expanding' when it was stated that space is expanding? Could the space of this room be said to be expanding? Authorities differed One would conclude that only the space between the nebulae was expanding, another, that the nebulae themselves were expanding as well The attempts to describe in terms of phenomena what was described by the mathematics almost always resulted in fantastic and extravagant conclusions; examples of such interpretations occur in Eddington's brilliant book, *The Expanding Universe*, wherein the universe is likened to an ever-widening running track, round which light ultimately began to describe its last lap, or to an inflating balloon, which ultimately bursts

Mathematicians have ever denied that the concepts they use can be really understood by such metaphors This is a reasonable position But I do not share the view that the real nature of phenomena is so obscure that they can only be described in terms of indefinable concepts which the plain man, not a mathematician, may for ever give up hope of understanding That might be so, but we can never establish it. We can but go on trying to gain insight into phenomena. The gaining of insight into a phenomenon is just as indefinable a notion as a mathematical concept But, as with a mathematical concept, we feel the better for it. It would be as wrong to preach that we can never 'understand' certain types of phenomena as to preach that we must never introduce a new indefinable concept My scientific prejudices in favour of clarity and against mysticism were brought to a head by a correspondence in *The Times*, in May 1932, on the subject of the curvature of space, a correspondence occasioned by Sir James Jeans's Ludwig Mond lecture at Manchester. One distinguished man of science was even reported as saying that 'we do not know, and probably shall never know, why space is expanding' This seemed to be carrying scientific pessimism too far, and it stimulated me to try to understand the expansion of the universe for myself The result was that from exceedingly elementary considerations it was possible to see how it came about that the universe was an expanding system, and why the relative velocities of the objects in it were proportional to their mutual distances The development of these ideas which had been prompted by a simple, common-sense view of the situation is the theme of this monograph The line of the development is obvious and inevitable, granted the initial idea, any

mathematician with sufficient faith in the usefulness and essential truthfulness of the initial idea would have travelled a similar path. Accordingly I do not feel personally responsible for the direction the investigations have taken, I felt merely that I was uncovering a situation already existing The investigations have taken many a surprising direction, and have at many a stage contradicted my intuitions and differed from my expectations, but always to prove themselves in thorough command of the phenomena The adventure has been more fascinating than I can say, and I can only wish to those who have the patience to read through the pages which follow even a fraction of the excitement and exhilaration which attended their working out

The principle of relativity

2. The outcome of the investigations is essentially independent of their application in astronomy. Formally the investigations may be said to have for their object the ascertaining of the motions of certain systems of particles without recourse to a specific theory of gravitation The motions obtained are such as must necessarily be followed by the particles enumerated in the description of the systems concerned. The motions are determined from kinematic considerations only, and are subject only to the condition of being determined by observation and being described in similar ways by observers, associated with particles of the system, whose rôles in relation to the system are indistinguishable The principle of relativity is in effect used in a new way, a way which is almost independent of observational verification, it is employed in the self-evident form that two observers, who stand in equivalent relations to the system and who agree to combine their observations (to yield coordinates) according to the same rules, will describe the behaviour of any particle by the same functions of these coordinates. The actual procedure will be clear from the later details

The principle of relativity is usually stated in the form that laws of nature are invariant in form for any arbitrary transformation of coordinates This is an exceedingly stringent requirement, imposing severe restrictions on the possible forms of laws of nature Here we impose the principle of relativity in a much weaker form, imposing far less severe restrictions and so leading to descriptions of permissible motions which are not apparently included, as far as is at

present known, within the totality of those permissible under the 'general' theory of relativity. The leading idea in our work is not that of transformations of *coordinates* but of transformations from *observer to 'equivalent' observer*, where the word equivalent will be strictly defined in terms of observations and tests which the observers can actually carry out. Transformations of coordinates alone are but translations of language, and have not necessarily much to do with phenomena. An observer can combine his observations to yield coordinates of events in an infinite variety of ways, coordinates are but arbitrary constructs out of observations. The important thing is not transformations of coordinates but transformations of observers, from one observer to another observer. A transformation of coordinates by a single observer, i.e. a new combination of observed data, leads to no new fact about the phenomena; it merely gives an alternative description of the phenomena by the same observer, that is, a new description of the same phenomena from the old point of view. We only get a new fact about the phenomena when we change the point of view, that is, when we change the observer. And we can only institute comparisons between the two points of view when the two observers agree to combine their observations (so as to yield coordinates) according to the same rules. When these rules have been agreed on to be the same, and when we know from the inner relationship of the observers that their points of view are similar, then we have a right to expect certain similarities in their descriptions. In particular, when the inner structure of the system defined is identical from the two points of view, then its description from the two points of view must be identical. This is the essence of the principle of relativity. It will be shown in detail to be the essence of the so-called 'special' theory of relativity, the 'general' theory of relativity introduces considerations and concepts which go beyond this. We confine attention to the coordinates which an observer can construct out of his own observations alone, made with his own measuring apparatus. The general theory of relativity, on the other hand, often employs coordinates which may be called 'mixed coordinates', coordinates constructed from observations made partly with the observer's own measuring apparatus and partly with the apparatus of other observers. An example is the 'cosmic time' t used in relativistic cosmology which is used as a coordinate and which is identical with a measure made by a second, distant, observer. We shall employ throughout

coordinates constructed out of observations made by an observer using his own measuring apparatus alone. We are strictly concerned with actual points of view, not hypothetical points of view whose description involves coordinates constructed out of 'mixed' observations

Whether this implies a criticism of the principle of relativity in its usual form must be left to the judgement of the reader The object of the present monograph is not to criticize the general form of the principle of relativity I rather wish to make clear that I throughout use a weaker form of the principle, and to claim that in their contexts results obtained by the application of this weakened form are unexceptionable It will be shown that this weakened form is adequate to define the motions we intend to discuss, which therefore depend for their validity only on a portion of the content of the principle of relativity in its stronger form Whether the stronger form is maintained or not, our present results must hold good But we see at once that we are entitled to an expectation that our results may include motions of systems excluded by the stronger form yet still perfectly valid

Thus it comes about that no appeal is made to 'general' relativity Still less is any appeal made to the Newtonian theory of gravitation or to Einstein's field equations of gravitation The methods are kinematic only, and no appeal is made to an assumed dynamics or to such concepts as force, mass, energy, momentum, which play no part. Lastly, no appeal is made to the 'special' theory of relativity in the form in which it is usually presented, though much of the mathematical analysis uses formulae identical with the Lorentz formulae and Einstein's velocity-addition formulae of 'special' relativity These formulae will be derived *de novo*, and will be used in contexts appropriate to their mode of derivation.

The method amounts to a new method of ascertaining the behaviour of particles in one another's presence in certain particular problems, and thus amounts to a new method for dealing with 'gravitation' Gravitation is simply a name for describing the motions that actually occur, motions which in the problems discussed are the only motions compatible with the principle of relativity in its weakened form. No causal influence is attached to gravitation, nor do we identify gravitation with an apparent field of force due to a geometry The geometry used by each observer is chosen arbitrarily *a priori*,

and not, as in the general theory of relativity, chosen to fit the distribution of given matter-in-motion present.

But not the slightest claim is made that the method amounts to a 'general' theory of gravitation. The theories of gravitation of Newton and Einstein are capable in principle of dealing with the motion of any gravitating system, whatever its composition. Here the method is only worked out for certain types of systems, namely systems containing 'equivalent' particles in the sense defined below. For these systems, the method prescribes all the details of the motion of every particle *given* to be present in the system, but it does not fully prescribe, in its present form, the motion of an additional free particle set loose in the system, that is, projected arbitrarily in the presence of the given particles forming the system. Whether it is open to generalization so as to make it capable of describing the motions of other systems, or of other free particles projected in the presence of other systems, will be discussed in general terms, but remains for future investigation. The position is therefore similar to that which would have arisen if the general solution of Kepler's problem (motion of a system consisting of one free particle and one 'heavy' particle) had been found from general considerations before the Newtonian dynamics and gravitation had been formulated,† or if similarly the two-body problem had been solved, or if Schwarzschild's form for the metric round a single point-mass had been found before Einstein's field equations were known, or again if the motion of a particular dynamical system, with a particular Lagrangian function, had been ascertained from general principles before Lagrange's equations of motion were known. I believe that the required generalizations are possible, and that a search for them would repay investigation.

3. The interest of the solutions might appear on this account to be distinctly limited. But it seems to me to be considerable for three reasons. First, the systems studied show that in certain cases motions which we conventionally ascribe to gravitation depend on no specific *formulation* of a law of gravitation at all, not even Einstein's, and therefore suggest that the same is true for any system, namely that all gravitational motions will be found to be the only possible ones compatible with the principle of relativity and the possibility

† As, in fact, it was.

of observing them, free from the restriction involved in the assumption that the world can be described by a Riemannian metric and in the assumption of specific 'field equations' with their present degree of arbitrariness. Secondly, the motions studied describe gravitating systems of which only limiting cases are described in Einstein's theory, it is a matter of distinct interest thus to isolate new solutions of the problem of gravitation, of a very simple character. Thirdly, the systems described and analysed by the method seem both suitable and adequate for a description of the astronomical universe, of which they reproduce many of the observed features, including several hitherto 'unexplained'.

The method involves no hypotheses or arbitrary assumptions. It introduces no empirical constants, when constants appear, it is merely as constants of integration or definite integrals, and they correspond simply to the degree of arbitrariness or generality in the systems discussed. The method simply sets about to enumerate the kinematic possibilities of the motions of the particles present, and it ultimately finds that the possibilities are so narrow that the motions are fully prescribed. It may be called the study of descriptive kinematics, for it merely describes, or rather enumerates observers' descriptions of, motions without appeal to any causal concept or laws of nature other than the principle of relativity in the above form. Though descriptions of gravitation in terms of action at a distance or in terms of the consequences of a modification of space due to the presence of matter are perfectly legitimate, they obscure the inevitability of the motions here disclosed. It is here claimed that this inevitability, this independence of any formulation of a *law* of gravitation, constitutes an advance, however small, on Einstein's theory.

4. After the method has been worked out without recourse to Einstein's theory, comparison will be made between our systems and certain systems described in modern relativistic cosmology. Other writers have already shown that in an 'expanding' space of constant negative curvature, as time goes on and the density of matter tends to zero, the observable phenomena tend to coincidence with a limiting form of one of our systems. The proof of this will be repeated using a different method. This method shows that the equivalence of the relativistic and kinematic systems depends primarily not on the curvature of the relativistic space, but on a certain dimensionless

parameter, equivalence arising when this parameter tends to infinity The latter limit is compatible not only with a limiting *zero* curvature for the relativistic space, but with a limiting *infinite* curvature, the density meanwhile remaining finite We thus obtain two distinct limiting cases of the relativistic systems and kinematic systems for which in the neighbourhood of the observer we get approximate coincidence But, observed at any finite epoch, the kinematic and 'general relativity' systems always differ in one fundamental respect The kinematic system always contains an infinity of particles in the field of view of any observer, merging towards the limit of visibility into a continuous background The systems of relativistic cosmology always contain a finite number of particles in the field of view of any observer, and these are seen against a non-luminous background, but as the epoch of observation advances, fresh particles continually appear in the field of view, increasing in number to infinity as the system tends to pass to approximate coincidence with the kinematic system Thus these relativistic systems imply *creation of matter* within the experience of any observer It will be found generally that all the systems proposed in general relativistic cosmology imply either the creation or the annihilation of matter within the experience of any observer, i e in time, which contradicts actual experience This characteristic therefore compels us to reject all the systems proposed in relativistic cosmology in favour of the kinematic systems, which always contain, at any epoch of observation, all their particles already in existence All the 'expanding space' solutions thus go, and the kinematic solutions alone survive

Space

5 A feature of the methods used is that the motions of the systems considered are described by means of the flat static space of ordinary physics, as remarked above, no recourse is made to a Riemannian metric, static or otherwise An advantage of this is that the systems can be pictured at once without the intervening calculations necessary when a Riemannian metric is used We almost unconsciously embed our ordinary eye-pictures in Euclidian space, and so our formulae convey eye-pictures When we make comparison with systems described by means of a Riemannian metric, we shall in all cases eliminate the geometry and compare simply the observational data that would be collected in the two cases by an assigned observer

Systems described by the methods of general relativity are expressed in terms of systems of coordinates which are of a conceptual character, never directly observed, and to picture such systems to the eye or the telescope or the photographic plate, calculations are required, sometimes of a complicated kind, to get from the coordinates used to the actual observations which they imply For this reason our treatment is essentially of a simpler character than the general relativity treatment

Throughout our work the standpoint adopted as regards space is that the space used for the description of phenomena is essentially arbitrary and at the disposal of the observer The phrase 'physical space' has no meaning The formulation of laws of nature, the description of phenomena in terms of coordinates, will be different according to the space selected, that is, according to the rules adopted for the conversion of observed data into coordinates Laws of nature and geometry are complementary, a modification of the one implies a modification of the other But it is always open to an observer to choose the flat static space of Euclid in which to embed the events the observations of which constitute his field of phenomena

This is an old view It is most clearly expressed, perhaps, in the writings of H Poincaré In *La Science et l'hypothèse*† he wrote 'La géometrie euclidienne n'a donc rien à craindre d'expériences nouvelles ' Again, 'Aucune expérience ne sera jamais en contradiction avec le postulatum d'Euclide' (he is alluding to the axiom of parallels), 'en revanche, aucune expérience ne sera jamais en contradiction avec le postulatum de Lobatchevsky', again, 'Il est donc impossible d'imaginer une expérience concrète qui puisse être interprétée dans le système euclidien et qui ne puisse pas l'être dans le système lobatchevskien ' The reason he gives is summed up in his aphorism 'Les expériences ont donc porté, non sur l'espace, mais sur les corps '

This view has tended to be abandoned owing to the successes of the general theory of relativity Space has been held to be 'really' curved near ponderable matter, the space has been held to be influenced by the presence of matter. For example, a distinguished student of relativity‡ has written 'The fundamental rule of general relativity is that in all models space, time, and material phenomena are interdependent; no one of them can be assigned without affecting the

† Chap V, L'expérience et la geometrie

‡ G C McVittie, *Observatory*, 57, 63, 1934

character of the others ' But this is a rule of procedure, not an assertion about some preconceived entity 'space' The field equations of relativity may in fact be regarded simply as a rule for selecting the space (calculating the $g_{\mu\nu}$'s which specify the space) given the matter present and its motion, the right-hand side of the equation

$$G_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}G+\lambda g_{\mu\nu} = -\kappa T_{\mu\nu}$$

specifies the matter-in-motion given to be present, whilst the left-hand side specifies the geometry.† General relativity moulds the geometry to the matter-in-motion, and chooses a different geometry for every different system It thus forges a weapon very powerful and very convenient for certain purposes, but not very convenient for others, not convenient, for example, for visualizing the phenomena in the space we habitually use in other branches of mathematical physics and in ordinary life But we must not argue that general relativity thereby denies the validity of the use of other spaces for other purposes. It is always within our choice to use a space different from the space appropriate to the description of the phenomena by the methods of general relativity, in particular, to describe phenomena in terms of, or by the use of, the space of ordinary physics I believe that this point of view has at present a considerable degree of acceptance For example, the alternative procedures have been described by S R Milner‡ as follows 'Two courses are open to us (1) We can modify the geometry assumed . . . so that a mathematically straight track (i e its length obeys a stationary principle) still continues to represent the non-uniform motion of a particle; this is the method of general relativity (2) We can retain the fourfold with unaltered geometry and specify a curved track which represents the observed motion by weighting each element of its length so that the integral weighted length between two points is stationary, this is the method of least action Both these methods of describing the motions of bodies must be considered equally logical, when one remembers that a manifold (even when it is called "space-time") is not the actual world, but a mental concept, in which real phenomena are represented

† Although, it may be remarked, the matter in-motion is assumed to be described in terms of the same coordinates as serve to map the geometry, so that a complete set of observations of the matter-in motion, by themselves, do not permit the immediate calculation of the right hand side The geometry has actually to be known first, before we can pass from observed data to the formulation of the right hand side Thus a conceptual element, namely the coordinates, enters explicitly at the outset

‡ *Proc Roy Soc*, **139**, 349, 1933.

symbolically ' I should prefer to call the manifold a 'mental construct'

The writings of de Sitter might persuade his readers into taking a different view Thus he writes † 'We have no intuitive knowledge of the kind of space we live in,' thus appearing to imply that we actually live in some particular kind of space, again, 'Euclidian space is a very close approximation to the true physical space ' He appears to believe that space has an inherent curvature, which, however, he thinks we can never determine 'We shall never be able to say anything about the curvature without introducing certain hypotheses '‡ But it must be remembered that de Sitter is talking about the geometry conventionally assigned for the description of the smoothed-out matter-in-motion of the universe by the general theory of relativity The actual geometry remains at the disposal of the observer, it is only when he voluntarily agrees to observe the rules of procedure of general relativity that de Sitter's statements are valid

Bertrand Russell has also expressed views|| which might be taken as affirming that there is a real entity, the space of nature Thus 'The physical problem may be stated as follows to find in the physical world, or to construct from physical materials, a space of one of the kinds enumerated by the logical treatment of geometry ' This is unexceptionable if he means the problem of actually carrying out the fit of an arbitrarily chosen space to a set of observed data defining phenomena But it is opposed to the view here advocated if he means the problem of finding *which* of the abstract geometries fits physical data And his continuation suggests that the latter view is the one he has in mind, for he writes later 'Many of the mathematically convenient properties of abstract logical spaces cannot be known to belong or known not to belong to the space of physics,' as though the space of physics was in fact one of these spaces, and, 'The subject of the physical theory of space is a very large one, hitherto little explored ' I should add that Russell has here in mind other properties

† *Kosmos*, p 117

‡ I believe that de Sitter is in error here It will be shown later that if we could make observations through long periods of time, then these observations would disclose whether the universe was following one of the model-careers predicted for it by relativistic cosmology, and if so what are the universal values of the conventional curvature and the cosmical constant Indeed, the same thing could be inferred by observations through a short period carried out with a sufficiently powerful telescope See Part IV

|| *Mysticism and Logic*

which may or may not be assigned to the space chosen to describe experience, such as continuity

Many more quotations might be given The above may serve as examples of positions held Here I wish only to establish a *prima facie* case, a case for the jury, in favour of the view that the observer may choose his space arbitrarily The test of this view is not in the last resort any academic discussion of its tenability The test is whether it works Can we put the point of view into practice, and describe by means of flat space systems containing matter of non-zero density for which general relativity would select a curved space? The investigations here gathered together do in fact put the abstract view to application in practice and describe certain systems of matter-in-motion in terms of the ordinary space used in physics It would be indeed remarkable if this were not possible, if the space which suffices for everyday experience and for the whole of physics excluding gravitation, were also inadequate for describing gravitational phenomena It would be no answer to this to say that space might 'really' have a *very small* curvature, for we should have to define a scale of smallness for the objection to have a meaning, and the system of nature would then have to contain a fundamental length relative to gravitational phenomena This could not be any 'radius of curvature' of the world, for that would be precisely the length we should be anxious to assess as large or small We shall in fact show that the world as a whole, in relation to the phenomena of matter-in-motion we discuss, contains in its description no fundamental length, but only dimensionless constants † Thus gravitational phenomena should be describable in terms of arbitrarily adopted flat or Euclidian space. It must be held to be a weakness of general relativity that its method obliges it to forgo this choice Clearly a need of present-day mathematical physics is a general method of handling gravitational problems in the space used in other physical problems This view is in violent opposition to the demand for a 'unified field theory', a demand for further geometrical modification of the space used so as to be able to employ these modifications in describing electromagnetic phenomena Both are possible modes of advance We can try to simplify laws of nature at the cost of complicating space, or simplify the space at the cost of complicating, or at least altering, the laws of nature

† When dynamical concepts are introduced, its description may be said to contain a fundamental mass as well

In denying any objective reality to any space used in physics, either that commonly used or those selected for the purposes of general relativity, I do not wish to imply that spatial relations have no meaning for us Our tactual experiences alone contradict this Further, we can construct spaces by the use of the indefinable concept of the rigid length-measure, and then fit such to experience by pretending that we recognize in certain objects exemplifications of rigid length-scales But no one has constructed a physics out of tactual data, and no one can be sure of having constructed a length-scale of unalterable length when transported, whatever that may mean Thus, to be very brief, spatial experiences exist and may be said to form part of the real world, but they are irrelevant to physics Russell has quoted Peano as remarking that the word *space* is one for which geometry has no use at all. Similarly physics has no use for spatial experiences We must endeavour to build up a physics out of observations without introducing the indefinable concept of the transport of rigid bodies

Time

6. This brings us to the subject of *time*, which seems to me to be on a totally different footing from space The passage of time is an undeniable constituent of our consciousness Of two events which happen to me, I can always say which occurs first, or whether they occur 'at the same time' Perhaps 'always' is too extreme a statement. But physics consists of over-extreme statements, and the idealized observer with which I identify myself is supposed to be able to carry out this judgement of assessing the 'earlier-or-later' relation between two events which occur to him Now if we exclude tactual experiences, and if further we exclude the use of the indefinable transportable rigid body, the coincidences of events which we call measures are always made *at the observer himself* They are events which occur at the observer The simplest such observation is the reading of a 'clock' at the moment the event occurs This is in principle simply the assigning of a number-sequence or number-continuum to events which occur in the observer's own experience The assigning of such a one-dimensional number continuum is possible just because of the above-mentioned immediate experience of a 'before and after' relation or a 'simultaneity' relation between events which occur to the observer himself We rely on the immediacy of judgement of the time sequence. Near the limit of simultaneity our judgement may

fail or may be ambiguous There, the treatment here to be adopted breaks down Until it so breaks down, that is to say, excluding too great 'nearness in time' of events that occur to me at myself, I can arbitrarily graduate an arbitrarily running clock, and read off from its dial the graduation or time-number of any event which occurs to me That is a basal judgement made on my experiences And all my experiences of the external world, excluding those mentioned above, are experiences at myself They constitute my outlook on the world, and their totality forms my data of observation The question arises whether such judgements, in the aggregate, form a sufficient basis for constructing a physics capable of dealing with the problems we propose to consider, that is, the problems of the motions of particles in one another's presence when such particles can be observed by me necessarily by observations recorded at myself.

Part I of this work is concerned with answering this question The answer is in the affirmative in so far as the ground we cover requires an answer The whole of our investigations, to which Part I is a necessary preliminary, rests on the immediate recognition of a temporal sequence amongst the events which occur to myself as observer †

We can make other judgements than recognitions of temporal sequence. We can recognize colours, match them, arrange them in a colour sequence We can recognize speeds; we can make an immediate judgement that one motor-car passes us at a faster rate than another (quite apart from watching whether the one overtakes the other) Doubtless a quantitative physics covering a certain domain of phenomena could be constructed out of such immediate judgements But no one has yet employed this course Here I am anxious merely to state a case for attempting to reduce the physics we need, considered as a set of measures, to measures based on temporal judgements.

Other judgements will be required to be introduced, but not, I think, other measures ‡ We shall need to make immediate judgements that two particles we observe are or are not in line with one another—coincide in apparent direction, or differ But the only measurements required will be measures made either with my clock or with some one else's

The subject of relativity only arises when we consider the experi-

† The treatment here given is essentially different from that in Dr A A Robb's *A Theory of Time and Space*, though starting from the same basis

‡ I refer to measures in terms of arbitrary standards Measures of solid angle are required

ences of a second observer If there were only one observer in the world, there would be no relativity For the one observer could transform his observations into coordinates in as many ways as he chose, and there would be no one to say him nay We require another observer, or indeed an army of other observers, whose experiences may be compared with one another, before there is any sense in speaking of a relativity, before a relativity can possibly exist Relativity and solipsism are incompatibles Relativity is the complete denial of the solipsist position Relativity is the comparison of experiences of different observers Thus it comes about that we introduce not only the particles we observe, but hypothetical observers on these particles outside ourselves There is no need to do so We could survey and describe the universe from ourselves alone, and let the matter rest there We should never contradict ourselves But we are not in fact content with descriptions of our own experiences alone, we desire to be able to describe what we should expect to see from other view-points, so we create, in imagination, other observers The other observers may simply be ourselves at some previous stage of our experiences, as when the observer in the Michelson-Morley experiment is carried by the earth round the sun, but he always implicitly assumes the possible hypothetical presence of another observer having experiences similar to those he himself formerly had The usual statements about changes of frames of reference postulate at bottom other observers located in these other, relatively moving, frames

We shall bring in these other observers, theoretically necessary in order that a subject of relativity shall exist, by explicitly attaching a hypothetical observer to other particles whenever we are interested in the point of view from these other particles Such other observers are to possess, like ourselves, temporal experiences We could not provide them with rigid length scales, copies of our own, because we could not say *a priori* what we meant by their being copies Nor can we provide them with clocks running and graduated in similar fashions to our own, because we cannot say *a priori* what we mean by 'similarly running' and 'similarly graduated' We can, however, suppose them furnished each with some kind of a clock, somehow graduated, because that is only equivalent to saying that each possesses a temporal experience It then becomes part of our task to state tests and procedures by which the different observers can compare their clocks, and come to some common agreement about their graduations, tests

by which a second observer can, if necessary, regraduate his clock to make it 'agree', in some sense yet to be fixed, with my own It will be shown that such tests can be stated in terms of the totality of the observations (time-measures) which the observers can make on one another using their clocks only When these tests have been made and the resulting calibrations effected, it will be found possible to say what we mean by a pair of observers in uniform or any other kind of relative motion, to derive the Lorentz formulae (or generalizations of them) connecting the coordinates they calculate out of their observations, and so to establish the formulae of the 'special' theory of relativity without using the indefinable concept of the rigid transportable length-measure That theory will be shown to depend on time-measures only, and to be simply the expression of the equivalence of particle-observers who find from their observations that they can describe one another as in uniform relative motion

It is to be particularly noted that we shall not assume as an axiom 'the constancy of the velocity of light' Since we have excluded the possession of rigid bodies, we can give no definition of velocity until we have agreed on certain combinations of time-observations as representing velocities, we cannot compare velocities so constructed until we have compared clocks, and by the stage at which we have succeeded in comparing 'relatively moving clocks' we have done all that the postulate of 'the constancy of the velocity of light' was invented to accomplish We finally prove that the number conventionally assigned by each of the 'equivalent' observers as his measure of the speed of light is the same for all observers

It is only with the meanings so attached to the famous Lorentz formulae and to 'special relativity' that we shall employ them in discussing gravitation and world-structure The 'space' we use will be throughout constructed out of time-measures Length coordinates, distance coordinates, will simply be certain combinations of time-measures

7. For our purposes, once we have constructed space out of time or temporal experience, nothing further is gained by fusing the observed time-measures and the constructed space-measures further into what is called 'space-time' The use of this new construct for mathematical purposes is at once available if we need it Actually, by way of illustrating the independence of our analysis of any belief in a 'higher

reality, space-time', I shall rarely if ever employ four-dimensional analysis even to perform mathematical calculations Still less shall I ever mystify my readers by using *imaginary* time, perfectly valid as that is as a geometrical concept What I sometimes lose in mathematical brevity† I hope to gain in clarity as to what each formula means in experience It is not space-time which is fundamental There is no absolute thing existing in itself which might be called 'the space-time of the world' Time-experiences alone are treated as fundamental We thus never encounter paradoxes, we never find the experienced time-sequence of the observer in apparent conflict with the 'relativity of time', as suggested in other treatments of relativity.

In his *New Background of Science*, Sir James Jeans admits that 'space means nothing apart from our perception of objects and time means nothing apart from our experience of events' This is in harmony with the above views But he regards what we here call 'constructed space-time' as 'objective space-time', and says ‡ 'Nature knows nothing of space and time separately, being concerned only with the four-dimensional continuum in which space and time are welded inseparably together into the product we may designate as "space-time" ' This is a most misleading statement To say that 'Nature knows ' is to make an assertion about nature for which we have no warrant, *we* do not know 'what nature knows', nor can we attach a meaning to the statement All we know are our observations of nature, fundamentally related to and expressed in terms of our temporal experiences, which we cannot deny Our temporal experiences are immediate, and we cannot put them on one side This immediacy of recognition of our temporal experiences is in flat contradiction with Jeans's next statement 'Our human spectacles divide this [i e "space-time"] into space and time, and introduce a spurious differentiation between them ' Why 'spurious'? This experience of ours of a temporal sequence is not spurious, it is, for us, one of the most real things in the world To make an artificial construct, space-time, out of our experiences, and then say that we 'separate the continuum into two ingredients' is to move in a circle No observer 'resolves his experiences into space and time', his temporal experience is something of which he has immediate awareness We shall be

† I will admit that I have often used the four-dimensional Minkowski 'picture' as a convenient way of first doing certain calculations

‡ p 101

content to put first things first, and base our treatment throughout on the observers' immediate awareness of the passage of time.

It will then appear that there is nothing 'objective', as Jeans maintains, about the speed of light. When two observers, reducing their temporal experiences of one another by the same rules, succeed in calibrating their clocks against one another, and evaluate what they choose to define as the speed of light using the same conventional definition, they arrive at the same number. The 'constancy of the speed of light' is a consequence of conventions.

To summarize, we begin with temporal experiences, construct distance-measures out of them, and combine the two, if we wish, into space-time. The only constituent of nature we have incorporated into our analysis is then the immediate recognition of temporal experiences. Relativity is the comparison of the experiences of different observers who can communicate with one another. The indefinable concept of the transport of rigid length-scales, or the 'slow transport of clocks',† is avoided.

Astronomical Applications

8. As stated earlier, the methods here used were originally developed as an attack on the subject of world-structure. They necessarily lead us to discuss problems of time, space, and relativity, and the problems of gravitation. The results, the gravitating systems we eventually describe, have an abstract interest independent of their applications. But their final interest lies in their application to the astronomical universe. This application is possible if it is assumed that the universe satisfies what is conveniently called 'Einstein's cosmological principle'. This is defined carefully in Chapter III. Here we are content to remark that it is simply a principle of selection, which chooses, out of all possible gravitating systems, a certain class as appropriate for representing the universe. The class selected contains no preferential particle as origin, and no unique standard of rest. The universe *might* be representable by a distribution of matter-in-motion which had a unique centre of symmetry, with the material distributed symmetrically about it, or it *might* have two foci, or it *might* be a rotating mass, or two tidally distorted aggregates, or any generalization of these. Whether it would be expected to be one of these, or whether it would be regarded rather as satisfying the cosmological principle, is a meta-

† Eddington's phrase.

physical question. Here we construct systems satisfying the cosmological principle and then use them simply as standards of comparison† with which to compare the observed universe.

The interest of these systems so constructed is then that they are found to reproduce many of the observable properties of the astronomical universe. They give at once the phenomenon of the expansion, and the expansion law (velocity proportional to distance). They give the aggregation of the material of the universe into sub-systems possessing a high degree of central condensation, and they predict the law of density-concentration, in fair agreement with observation. They give the form of the trajectories of the constituents of the sub-systems. They predict the presence, in any given volume of space, of particles moving with speeds arbitrarily close to that of light, and show how they have attained these speeds by falling towards the apparent centre of the universe reckoned in the frame in which they are momentarily at rest. Lastly, they predict the presence, near the centre of any given sub-system, of particles which have come nearly to rest relative to this sub-system—particles which have arrived in this vicinity from other, distant, sub-systems to which they originally belonged. The swiftly moving particles may provisionally be identified as the primary agents producing the well-known cosmic rays, for their impacts with other particles may be expected to give rise to the corpuscular radiations observed in ionization chambers. The slowly moving particles, which ultimately come to rest relative to the sub-system concerned, may be provisionally identified with the 'cosmic clouds' or other obscuring clouds which appear to occupy the vicinity of our own galaxy and which are seen as bands of obscuring matter in the diametral planes of spiral nebulae. Whether these identifications

† In earlier writings on this subject I used the cosmological principle as a weapon of investigation, i.e. assuming that the universe *obeyed* it I attempted to infer the properties it would have. In the present monograph the cosmological principle is regarded throughout simply as a definition, defining what system we propose to consider. In any gravitational problem, we have to define the systems under consideration; for example, if we wish to consider the abstract two-body problem, we confine attention to solutions possessing just a pair of singularities; similarly, if we wish to consider rotation of an incompressible mass, we define our system as one possessing a definite density, definite angular velocity, and definite boundary conditions. The cosmological principle is simply used as a definition, a principle of exclusion, enumerating the class of systems to be considered and excluding all others. It is in no sense used as a 'law of nature', or principle of compulsion, prescribing what is to happen. Whether, when a system is set up satisfying the cosmological principle, it will continue to satisfy it, is a matter always for investigation. I hope that this explanation meets Professor Dingle's published criticisms.

M 101, N G C 5457, *Ursa Major*, Spiral Nebula, exposure 7 hours 30 minutes, March 10 and 11, 1910 60 inch Reflector

(By the courtesy of the Director of the Mount Wilson Observatory)

prove justifiable or not, the presence in any region of space of swift particles and near galaxies of slow particles would seem to hold good on any theory of gravitation

The present methods give no express account of the rotations of galaxies or of their spiral forms But they leave a natural place for extensions to these The methods of statistical kinematics represent the systems of particles described as quasi-continuous dust-clouds, and so bear the same relation to descriptions of discrete, finitely separated particles as the differential calculus bears to the calculus of finite differences They give *average* effects only They make no allowance for the *discreteness* of the particles and sub-systems concerned Certain features exhibited by quasi-continuous dust-clouds appear not to be reproducible by systems of particles with finite separations, the finite separations seem to forbid the possibility of satisfying the cosmological principle in its strict form The resulting departures from perfect equivalence of particle to particle will involve local variations in the aspect of the system from particle to particle, and such variations of aspect may be expected to be associated with departures of the values of the acceleration from the values obtained by statistical methods, these in turn will imply locally non-homogeneous fields, tidal couples and the like (interpreted dynamically), and to these, following Jeans, we may attribute the forms and internal motions of the nebulae other than those accounted for by the statistical theory The problem here seems bound up with the three-dimensionality of space, finite-difference solutions that are possible in one dimension seem to have no analogues in three dimensions

The investigations which follow supersede those published in any earlier papers, and follow a different logical sequence

PART I

KINEMATICS AND RELATIVITY

II

EQUIVALENT PARTICLE-OBSERVERS

9. THE present chapter is concerned with the fundamental ideas of relativity. It proceeds without making use of the experimental facts on which the 'special' or 'restricted' theory of relativity is usually based, though its results coincide in large part with the formal relationships developed in the special theory of relativity, here the results are developed in new contexts, suitable for immediate application to problems of world-structure Again, it proceeds without recourse to the conceptual methods of the 'general' theory of relativity, the assumption is nowhere made that phenomena or events can be located by means of conceptual coordinates in a Riemannian metric In general relativity, coordinates are introduced at the start as a possible conceptual mode of description of events, and are then translated into the possible observations which they imply, here, we shall begin with observations, and any coordinates used will be constructs out of observations that have already been made by the relevant observers, or that can in principle be made by them The object of the investigation is thus to construct a calculus on the basis of observations that could actually be carried out, without the use of indefinable concepts We shall not employ the concept of the transport of invariable or rigid length-scales, because this is an indefinable concept. Without considerable analysis it is impossible to say what is meant by asserting that a given measuring-rod and another given measuring-rod in another place or in motion relative to the first have the same 'length', and it is unwise to begin a fundamental investigation by introducing this notion as an indefinable concept Nor shall we employ the concept of the transport of invariable clocks, it is not possible without careful analysis to say what is meant by the assertion that two clocks in different places or in relative motion keep the 'same time' As a consequence of this act of self-denial, we shall not be in a position to employ the axiom of 'the constancy of the velocity of light to observers in uniform relative motion'; for

velocity is the ratio of a length to a time-difference, and neither will as yet have been defined Instead we shall state tests, which could actually be carried out, by which observers in different places and in relative motion could compare their clocks and measuring-rods, and we shall find that when they adopt the same arbitrary conventions for calculating numbers which they call coordinates out of directly observed numbers observed by means of properly compared clocks, the numbers they are led to assign to what they agree to call the velocity of light come out, under certain circumstances, the same

It seems to me that here our investigations break fresh ground In his admirable Gifford lectures, Dr Barnes has quoted with approval an aphorism of Poincaré 'If there were no solid bodies we should not have geometry' Yet it is remarkable that relativistic cosmology describes as systems suitable for representing the world not systems of solid bodies but systems of particles, the nebulae are idealized as a dust-cloud of particles The rigid body, in these presentations, comes in in the foundation of the theory, but plays no direct part in the systems of particles in motion they eventually describe In discussing particles only it should be possible to proceed on a particle basis alone The rigid body plays indeed very different rôles in different parts of these theories, sometimes as a rigid length-unit, sometimes as a theodolite or rigid angle-measurer These two usages are to be sharply distinguished, for the metre is an incommunicable unit, primarily, whilst it is possible to communicate to a distant geometer what is meant by an angle of 60° We shall at times employ the notion of the existence of the theodolite, but not the notion of the metre-scale. In so far as our procedure is free from criticism, this will prove some advance towards the ideal of describing systems of particles with recourse to particles only.

10. We shall employ the concept of 'particle', and associate with each particle an 'observer' The combination will be called a 'particle-observer' The particle can be the seat of events, and these events can be observed by the associated observer The observer is supposed to be capable of describing his observations to other observers and to be capable of dispatching, at any given event which occurs at himself, a signal to other observers at the moment of occurrence of the event at himself

11. In applications of relativity to world-structure, a considerable part is played by particles which are equivalent to one another in the sense that a given particle-observer A describes his observations on surrounding particles in the same terms as any other given particle-observer B describes *his* observations on *his* surroundings. For this to have a meaning it is necessary for A and B to agree on codes of description of the observations of the events E which occur at themselves, the events E being the receptions of signals from other particles. Further, we by no means wish to confine consideration to cases in which every particle is equivalent to every other. It therefore seems best to consider first the case of a pair of particle-observers who are equivalent to one another when attention is paid only to the observations they can make on one another, i.e. excluding observations on other particles.

We shall therefore say that two particle-observers A and B are *equivalent* when the totality of observations A can make on B can be described by A in the same way as the totality of observations which B can make on A can be described by B. When this condition is satisfied, we shall write $A \equiv B$. A system of particles is then said to satisfy Einstein's cosmological principle when if A and B are two members of the system such that $A \equiv B$, then A's description of the whole system is identical with B's description of the whole system.† Whether the universe, idealized to a system of particles, may be expected to satisfy Einstein's cosmological principle is partly a metaphysical question, whether it does in fact satisfy it is a question of experiment and observation, a question of test of fact, whether systems can be constructed satisfying the cosmological principle is a question of mathematics. We shall later examine these questions. But before we construct systems of particles satisfying Einstein's cosmological principle, it appears to be an indispensable preliminary to investigate the conditions under which A is equivalent to B in the narrow sense defined above, without consideration of other particles (other than A and B alone) which may be present. For the moment we are content to remark that in a system satisfying Einstein's cosmological principle, it is by no means necessary that every pair of particles should be equivalent to one another, the system, described

† Einstein originally stated his principle in the form 'Alle Stellen des Universums sind gleichwertig'. Cf. *Berlin Sitz.*, 235, 1931. This is a very vague statement, but it justifies the coupling of Einstein's name with the more precise form of the principle enunciated above.

in the same way by A and by B, may contain and will in general contain, other particle-observers who will describe the system in different terms; though it may contain also any number of particle-observers equivalent to A and to B In the present chapter we confine attention in the first instance to the equivalence of specified particle-observers A, B, C,

Clocks

12. Our definition of equivalence is still not complete The phrase 'the totality of observations A can make on B' implies the possession by A and B of measuring apparatus In astronomical practice this apparatus consists essentially of clocks, telescopes, cameras, spectroscopes, and photometers, an astronomer can time, count stars, measure their spectra and their brightnesses, and ascertain their apparent distribution or relative proximity in his field of view In an abstract investigation dealing with fundamentals it is necessary, however, to reduce this collection of apparatus to its essentials The more abstract an investigation is at the outset, the more concrete it can be made later. We shall therefore inquire what progress can be made by using *clocks* only. It is in any case necessary to confine attention to a single piece of measuring apparatus For two observers might be equivalent in regard to one piece of apparatus and not in regard to others.

We shall therefore suppose each observer A and B armed with a clock We must say what we mean by a clock, and describe how it is to be graduated.

13 Each observer is supposed to possess a temporal experience. That is to say, of any two events E_1 and E_2 which occur to himself, or at himself, he can make an immediate judgement as to whether E_2 follows E_1, precedes E_1, or is simultaneous with E_1 Further, this temporal experience is supposed to be continuous: between any two non-simultaneous events E_1 and E_2 other events can occur at A, and be judged as preceding E_2 and following E_1, if E_2 is later than E_1 The events which constitute A's temporal experience will, still further, be supposed to constitute a continuum That means that A can correlate the totality of events in his temporal experience with the real numbers Associated with any event E which occurs to A is a number t, which we call the epoch of E Such a correlation of real numbers with the events constituting A's temporal experience we call a *clock*. It may be pictured as a hand running continuously round a dial, no

matter how 'irregularly', the dial being graduated in some fashion and provided with a revolution-counter When some event E occurs at A, A may prick off the simultaneous position of the hand, by an immediate judgement, and read off the epoch t of E

B is supposed also to possess a temporal experience, and to carry a clock. The question arises how is B to graduate his clock so that he can attach a meaning to saying that he is in possession of a clock identical with A's? We shall solve this problem for the case $A \equiv B$ under the further restriction that at some one epoch A and B parted company. That is to say, we shall state tests by which A and B can ascertain whether they are equivalent, tests which they could actually carry out for various clock graduations

14. The restriction that A and B once parted company is not an essential restriction, it is made in order to afford an easy method for A and B to synchronize the zeros of their clocks They are to agree that the moment at which they parted company is to be labelled '$t = 0$' If in their experience the two particles never coincided, the synchronization of zeros, which may be carried out in a variety of ways, by selecting specific events in their experiences as $t = 0$, is arbitrary I have not had the leisure to carry through the analysis when A and B pursue non-intersecting trajectories, and this problem would certainly repay investigation. Not only so, but it would have far-reaching cosmological applications. Our object is ultimately to construct systems of particles containing equivalent pairs, and to gain insight into the structure of such systems it seemed best to deal first with the case of intersecting trajectories As regards non-intersecting trajectories, the simplest case is that of 'relatively stationary particle-observers', but when we dispense with the concept of the rigid length-measure it is not at once apparent what is meant by being 'relatively stationary'. I have in fact analysed this 'stationary' problem, with surprising results, but they will not be treated here, they are sufficient to show what a rich field awaits investigation, even in that simple everyday case of determination of longitude differences by observers provided with clocks and wireless-signalling apparatus Here, then, we confine attention to intersecting experiences

Possible observations

15. One of the main features of the investigation which follows is the recognition that if A and B are armed with clocks only, the

totality of observations that A can make on B can be reduced to two types.

(1) A can send a signal to B at time t_1 by his (A's) clock, and arrange with B that it is returned by B (or reflected by B) at the moment of reception of the signal by B, A can then observe, by his own clock, the time of reception of the echo-signal from B, say t_2. The epochs t_1 and t_2 are numbers observed by A with his own arbitrarily graduated clock. A can repeat these observations as often as he chooses (using some method such as the Morse code or coloured lights to identify return signals). The determination of t_2 as a function of t_1 determines A's description of B's motion 'in the line of sight'.

(2) A can read B's clock directly at the epoch t_2 by his (A's) own clock. We can if we like give A a telescope to carry out this operation, A simply reads B's clock at the moment of return of the echo-signal. But it is sufficient if B transmits, with his return signal, the epoch t'_B by his (B's) clock at which the direct signal was received and returned. Alternatively B can transmit this information at his leisure. The determination of the epoch t'_B as a function of t_1 or of t_2 fixes the behaviour of B's clock in A's experience.

Though such interchange of signals may be considered as far removed from astronomical procedure, it is in fact closely connected with it. The Doppler effect for B observed by A will be found to be simply the differential coefficient dt_2/dt'_B, once the clocks have been graduated so as to be comparable. The whole of the present analysis can in fact be carried out by starting with Doppler effect observations and subsequently integrating them, but the observational procedure here outlined, though more abstract, is more primitive.

We shall find that the whole problem of the descriptions of *any* event (not necessarily at A or B) by A and by B turns on the determination by A of the running of B's clock in A's experience.

B can carry out similar observations on A. He can send a signal at time t'_3 by his clock, receive the echo-signal at time t'_4, and ascertain the reading t_A of A's clock at the moment of reflection by A, by an immediate judgement by A. He can then plot t'_4 and t_A as functions of t'_3.

A and B will then be equivalent when t_2 is the same function of t_1 as t'_4 is of t'_3, and when at the same time t'_B is the same function of t_1 or t_2 as t_A is of t'_3 or t'_4.

We shall show here that consistency of observations implies that

the second function is determinate when the first is prescribed The reason that the functions are interconnected is that the same signal may be regarded as an outward signal from A or as a return signal to B The form of the second function is known from the special theory of relativity when A and B describe themselves as in 'uniform relative motion', a phrase to which we have as yet given no meaning Here we shall not confine ourselves to 'uniform relative motion', but deal with the arbitrary relative motion of equivalent observers We shall thus be led to a generalization of the special theory of relativity.†

Coordinates

16. It is possible to conduct the following analysis in terms simply of the observational data t_1, t_2, t'_B, t'_3, t'_4, t_A But its relation with ordinary physics is better displayed if we now introduce coordinates Coordinates are numbers constructed out of observational data

Particle-observer A has three data of observation associated with the event E_B which was the event of the reflection of a certain signal by B Of these three, t_1 and t_2 are readings on A's clock, t'_B is a reading on B's clock The description A formulates of the behaviour of B in A's experience should clearly depend on the apparatus used by A for observing B's behaviour, and be independent of the apparatus that happens to be in the possession of B We therefore impose the requirement that coordinates A assigns to event E_B shall depend on the measures t_1, t_2 made by A with A's clock only, and shall not depend on t'_B, the reading of B's clock (In general relativity t'_B is frequently used as a coordinate, and called the 'cosmic time' of the event E_B, but it cannot be immediately determined by A by measures with his own apparatus General relativity uses the word 'coordinate'

† The present method suggested itself after I had proposed in lectures at Oxford a possible mode of derivation of the Lorentz formulae by idealized experiments involving light signals only, assuming the concept of uniform relative motion and the postulate of the constancy of the velocity of light to two particle-observers in uniform relative motion The derivation was effected by Mr G J Whitrow in a paper published in *Quart Journ Math* (Oxford), 1933 I then found it possible to simplify the ingenious but somewhat complicated procedure due to Whitrow, and to reduce the problem of the proof of the Lorentz formulae to the determination of the running of one observer's clock in the experience of the other observer It then appeared that there was no necessity for the restriction to uniform relative motion The investigations forming the present chapter were delivered as lectures at seminars in Oxford in the October term, 1933 The relations of the methods here developed to the classical methods of derivation of the Lorentz formulae are not discussed here, they have been discussed by Whitrow in the paper cited, with full references It is scarcely necessary to remark that the present account does little more than open up the subject

to represent any combination of possible measures, whether made by one observer or compounded out of observations made by more than one observer We shall find that considerable clarity results from confining the word coordinate to constructs from 'pure' measures by a single observer and excluding constructs from 'mixed' measures)

Out of the two observed numbers t_1 and t_2 determined by immediate judgements by A using his own clock, A can construct two other independent numbers in an infinite variety of ways We desire to construct two numbers one of which we may call the *epoch* of E_B by A's clock, the other the *distance* of E_B by A's clock The epoch we desire to assign to E_B must for consistency be a number exceeding t_1 and less than t_2, for in any sense in which we can attach a meaning to 'earlier' and 'later' E_B is later than the event at A of epoch t_1, earlier than the event at A of epoch t_2 Further, if we renumber A's clock-graduations by the addition of a constant t_0, we require the *epoch* assigned to E_B to be increased by the same constant, and we require the *distance* assigned to E_B to be independent of the addition of a constant to A's clock-graduations Lastly, we require that when E_B occurs at A, the *epoch* of E_B shall reduce to the epoch by A's clock, and the *distance* shall reduce to zero It is readily shown that these conditions imply that the epoch-coordinate of E_B must be of the form

$$\tfrac{1}{2}(t_2+t_1)+\psi_1(t_2-t_1),$$

and that the distance-coordinate of E_B must be of the form

$$\psi_2(t_2-t_1),$$

where $\psi_1(0) = 0 = \psi_2(0)$ The simplest choice for ψ_1 and ψ_2 is to take $\psi_1(t_2-t_1) \equiv 0$, $\psi_2(t_2-t_1) \equiv \tfrac{1}{2}c(t_2-t_1)$, where c is a positive number arbitrarily chosen by A We shall now write

$$T_B = \tfrac{1}{2}(t_2+t_1), \qquad R_B = \tfrac{1}{2}c(t_2-t_1) \tag{1}$$

and call T_B the epoch of E_B, R_B its distance Since $t_2 > t_1$, necessarily $t_2 > T_B > t_1$ and $R_B > 0$ It should be noted that T_B and R_B are purely conventional constructs In this chapter, small letters t_2, t_1, etc , denote the results of immediate judgements, whilst capitals T_B and R_B denote conventional coordinates

In order to compare his observations with A's, B must adopt the same rules as A for constructing his coordinates of events E_A at A out of his observations by his clock By agreement with A, he chooses the same positive number c and writes

$$T'_A = \tfrac{1}{2}(t'_4+t'_3), \qquad R'_A = \tfrac{1}{2}c(t'_4-t'_3) \tag{1'}$$

B could quite well choose a different number for c But A informs B of his choice, and B elects to make the same choice

17. From repeated observations of events E_B at B, A can determine R_B as a function of T_B, and also t'_B as a function of T_B (A simply draws graphs by plotting R_B and t'_B against T_B) Let these functions be $R_B = c\phi_{12}(T_B)$, $t'_B = f_{12}(T_B)$ Similarly B can determine R'_A as a function of T'_A, and t_A as a function of T'_A Let these functions be $R'_A = c\phi_{21}(T'_A)$, $t_A = f_{21}(T'_A)$ The assignment of distances and epochs are shown diagrammatically in Fig 1, the arrows indicating the observer making the assignments

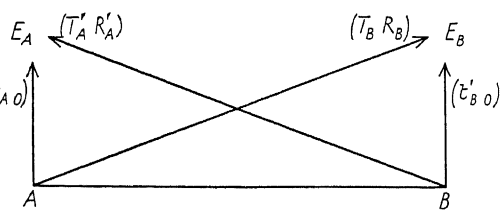


FIG 1 Observations and calculations by equivalent particle-observers on one another

A and B can now try different graduations for their clocks, and for each graduation-scheme of either clock compare the functions ϕ_{12} and ϕ_{21}, and the functions f_{12} and f_{21} A and B will be *equivalent* if clock-graduations exist such that their graphs are identical, i e such that

$$\phi_{12} \equiv \phi_{21}, \equiv \phi, \qquad f_{12} \equiv f_{21}, \equiv f,$$

so that

$$R_B = c\phi(T_B), \quad R'_A = c\phi(T'_A), \quad t'_B = f(T_B), \quad t_A = f(T'_A) \qquad (2)$$

The test of equivalence is thus one that can be actually carried out Every alteration of clock-graduations alters the conventional coordinates T_B, R_B, T'_A, R'_A and so alters the various graphs and the functions which describe them

We shall now assume that it is possible to realize in nature 'equivalent observers', that is to say that pairs of observers could be set up which would move (or could be moved if necessary by appropriate constraints) such that $\phi_{12} \equiv \phi_{21}$, $f_{12} \equiv f_{21}$ This is the experimental basis of our investigation The situation contemplated must be ideally possible But the investigation can also be regarded as an abstract geometry The propositions which follow are true of entities which are the subjects of the propositions $\phi_{12} \equiv \phi_{21}$, $f_{12} \equiv f_{21}$ inde-

pendently of whether such entities 'exist' or not, just as propositions of geometry are true of the subjects of certain axioms without the 'existence' of these subjects being relevant Actually if it were not possible to realize in nature equivalent observers the subject of relativity as a branch of experimental philosophy would not exist

Conditions of equivalence

18. We proceed to show that if $\phi_{12} \equiv \phi_{21} \equiv \phi$, $f_{12} \equiv f_{21} \equiv f$, then a relation exists between ϕ and f In words, this means that when two particle-observers are equivalent, their relative motion determines the behaviour of either clock in the experience of the other

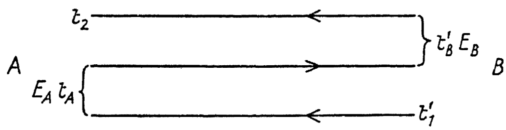


FIG 2 Scheme of interlocked signals

observer When two observers are equivalent, we can say that they are provided with *identical* clocks, where a test of identity is implied which could actually be carried out

For consider the scheme of signals depicted in Fig 2 Let B dispatch a signal at epoch t'_1 by his clock, and let A reflect it at epoch t_A by A's clock, call this event E_A Let B receive this reflected signal at t'_B by B's clock, and call this event E_B Lastly let A receive the signal reflected at E_B at time t_2 by A's clock This sequence of signals can be repeated as often as we like, and t'_1 may be chosen arbitrarily The four times t'_1, t_A, t'_B, t_2 are immediate judgements by the observers

A and B now assign coordinates (epochs and distances) to events E_B and E_A as follows·

Event E_B A's assignments

$$T_B = \tfrac{1}{2}(t_2+t_A), \qquad R_B = \tfrac{1}{2}c(t_2-t_A) \tag{3}$$

Event E_A B's assignments

$$T'_A = \tfrac{1}{2}(t'_B+t'_1), \qquad R'_A = \tfrac{1}{2}c(t'_B-t'_1) \tag{4}$$

From these equations, eliminating t_2 and t'_1, we have

$$T_B-R_B/c = t_A,$$
$$T'_A+R'_A/c = t'_B$$

But

$$R_B = c\phi(T_B), \qquad R'_A = c\phi(T'_A),$$

$$t'_B = f(T_B), \qquad t_A = f(T'_A)$$

Hence

$$T_B - \phi(T_B) = f(T'_A),$$

$$T'_A + \phi(T'_A) = f(T_B)$$

Since T_B, T'_A depend on one arbitrary variable t'_1, it follows that if T'_A is chosen arbitrarily and T_B found from the first equation and substituted in the second, the second must be an identity in T'_A

More briefly we may say that the relations may be written, on putting $x = T_B$, $y = T'_A$,

$$f(y) = x - \phi(x), \tag{5}$$

$$f(x) = y + \phi(y), \tag{6}$$

and these equations are such that if the value of either x or y is found by solution of one of the equations, in terms of y or x respectively, and substituted in the other, then the resulting equation must be an identity in the surviving variable Equations (5) and (6) will be found to imply the whole of 'special relativity' and a great deal more.

Our problem is to determine f, given ϕ These equations reduce to a question of mathematics the problem of what B's clock reads in the experience of A, given that A and B are equivalent and that A has ascertained by observation B's distance-time relation ϕ in terms of his own (A's) coordinates

19. Let us examine the circumstances in which $f(x) \equiv x$ is a solution, i e the circumstances in which B's clock at any event at B reads the same time as the epoch conventionally assigned to the same event by A (in this case A's and B's clocks might be said to 'agree') This requires that

$$y = x - \phi(x), \qquad x = y + \phi(y)$$

should give an identity when x or y is eliminated Adding, we have $\phi(x) = \phi(y)$. If the 'motion' of B in A's experience is progressive, ϕ is one-valued, and we have either $x = y$ or $\phi \equiv$ constant If $y = x$, then $\phi(x) = 0 = \phi(y)$ and the two particles are permanently coincident If $\phi \equiv$ constant, the two observers reckon one another as at a fixed distance apart, and thus at relative rest † It follows that when

† In this case they have never coincided, and our assignment of a zero to the clock-readings requires reconsideration It is easily seen, however, that in this case the observational relations are now independent of the addition of an arbitrary constant to the readings of either clock

relative motion occurs ($\phi \not\equiv$ const), $f(x)$ cannot be represented by x, and so $T'_A \neq t_A$, $T_B \neq t'_B$ This is Einstein's fundamental discovery

It is convenient to say that all events to which A attaches the same epoch T form 'the world-wide instant $t = T$' for A Thus in A's world-wide instant $t = T_B$, B's clock does not read T_B. If we call B's time for an event at himself, namely t'_B, the 'cosmic time' of that event, then the cosmic time of an event never agrees with A's coordinate time for the same event when A considers B as in motion relative to himself

Solution in terms of a parameter function

20 We now obtain a parametric solution of equations (5) and (6) Let f^{-1} denote the function inverse to f, this is easily seen to exist Define functions p_{12}, p_{21} by the relations†

$$p_{12}(\xi) = f^{-1}(\xi) - \phi f^{-1}(\xi), \tag{7}$$

$$p_{21}(\xi) = f^{-1}(\xi) + \phi f^{-1}(\xi) \tag{8}$$

Replacing ξ by $f(\xi)$, we see that these imply

$$p_{12} f(\xi) = \xi - \phi(\xi), \tag{7'}$$

$$p_{21} f(\xi) = \xi + \phi(\xi) \tag{8'}$$

Hence (5) and (6) may be written

$$f(y) = p_{12} f(x), \tag{9}$$

$$f(x) = p_{21} f(y) \tag{10}$$

Eliminating y, we have $p_{12} f(x) = p_{21}^{-1} f(x)$

which must be an identity in x Hence p_{12} and p_{21} are inverse functions, they may be considered as inverse operators

This leads to the following rule for constructing solutions of (5) and (6) in terms of a parameter function p_{12} Take an arbitrary function $p_{12}(\xi)$, defined for all $\xi > 0$, and possessing an inverse $p_{21}(\xi)$ Define functions $f(\xi)$ and $\phi(\xi)$ by the relations

$$f^{-1}(\xi) = \tfrac{1}{2}[p_{12}(\xi) + p_{21}(\xi)], \tag{11}$$

$$\phi(\xi) = \tfrac{1}{2}[p_{21} f(\xi) - p_{12} f(\xi)] \tag{12}$$

Then the functions f and ϕ so constructed satisfy (5) *and* (6) For, replacing ξ by $f(\xi)$ in (11), we have

$$\xi = \tfrac{1}{2}[p_{12} f(\xi) + p_{21} f(\xi)],$$

whence, using (12),

$$\xi - \phi(\xi) = p_{12} f(\xi),$$

$$\xi + \phi(\xi) = p_{21} f(\xi)$$

† The notation $\phi f^{-1}(\xi)$ means $\phi(f^{-1}(\xi))$, and similarly for all other functional operators

It follows that (5) and (6) will be satisfied provided the equations

$$f(y) = p_{12}f(x), \qquad f(x) = p_{21}f(y)$$

yield an identity in either of the variables when the other is eliminated This is satisfied since p_{12} and p_{21} are inverse functions The solution x, y of (5) and (6) may now be put in either of the forms

$$y = f^{-1}p_{12}f(x), \qquad x = f^{-1}p_{21}f(y) \tag{13}$$

Moreover, (11) and (12) provide the most general solution of (5) and (6) For if f and ϕ exist, so do p_{12} and p_{21} by (7) and (8), and by (9) and (10) these are inverse functions The functions p_{12} and p_{21} play a fundamental part in what follows

Doppler effects

21. The physical meanings of the functions p_{12} and p_{21} are readily found. We have seen that the observers A and B can make definite tests of their equivalence in terms of their clock-readings Equations (5) and (6) then show that the clock-behaviour function f is determinate (or at least given by a specific functional equation) when ϕ is known The time-distance function ϕ determines the clock-behaviour function f prescribing the behaviour of 'identical' clocks in the experience of the relatively-moving observers carrying them This affords an objective meaning to the carrying of 'identical' clocks by equivalent observers We may, however, if we choose, suppose that we can make an immediate recognition of whether two clocks are identical or not The simplest ideal case where such a recognition is possible is the case of the atomic clock—two 'identical' atoms carried by the observers Then our analysis at once gives the apparent relative running of the two clocks, in terms of the relative motion, and this is easily seen to be equivalent to an evaluation of the Doppler effect

For consider the simple signals depicted in Fig 3

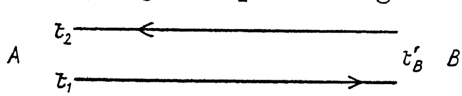


FIG 3 Direct and return signal made by A

With our previous notation we have

$$\tfrac{1}{2}(t_2+t_1) = T_B, \qquad \tfrac{1}{2}c(t_2-t_1) = R_B,$$

where $$R_B = c\phi(T_B).$$

Hence $$T_B+\phi(T_B) = t_2$$

Also $$t'_B = f(T_B)$$

We want to eliminate T_B between these equations From the properties of the p_{12} function already given

$$T_B+\phi(T_B) = p_{21}f(T_B).$$

Hence† $$t_2 = p_{21}(t'_B) \quad \text{or} \quad t'_B = p_{12}(t_2)$$

Now a signal leaving B at t'_B by B's clock arrives at A at t_2 by A's clock, and a signal leaving B at $t'_B+dt'_B$ by B's clock arrives at A at t_2+dt_2 by A's clock. Hence the signal-interval dt'_B recorded on an atomic clock at B is received at A as an interval dt_2 If ν_0 is the frequency of the light emitted by B, ν'_0 that received by A, then when $dt'_B = 1/\nu_0$, we must have $dt_2 = 1/\nu'_0$ Thus if λ_0, λ'_0 are the wavelengths corresponding to ν_0 and ν'_0,

$$\left(\frac{\lambda'_0}{\lambda_0}\right)_A = \frac{\nu_0}{\nu'_0} = \frac{dt_2}{dt'_B} = p'_{21}(t'_B) = \frac{1}{p'_{12}(t_2)}$$

Thus the differential coefficient $p'_{21}(t'_B)$ measures the Doppler displacement for the observation made by A at the epoch of observation $t_2 = p_{21}(t'_B)$ by A's clock It follows that the differential coefficient $p'_{21}(t'_B)$ is directly observable by A If we denote by $s(t_2)$ the Doppler shift-ratio observed by A at epoch t_2 by A's clock, then

$$s(t_2) = p'_{21}(t'_B), \qquad dt_2 = p'_{21}(t'_B)dt'_B,$$

so that $$dt'_B = \frac{dt_2}{s(t_2)}.$$

Consequently, apart from an additive constant, t'_B can be found by integration of actual observations made at A, namely by integration of $dt_2/s(t_2)$, given that A and B are equivalent.

Similarly B can determine A's local time by observations of the Doppler shift observed at B For from the scheme of signals shown

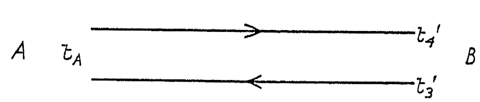


FIG 4 Direct and return signal made by B

in Fig 4 it follows by an identical calculation that

$$t'_4 = p_{21}(t_A) \quad \text{or} \quad t_A = p_{12}(t'_4),$$

whence $$\left(\frac{\lambda'_0}{\lambda_0}\right)_B = \frac{dt'_4}{dt_A} = p'_{21}(t_A) = \frac{1}{p'_{12}(t'_4)}$$

When t'_4 happens to equal t_2, $(\lambda'_0/\lambda_0)_B = (\lambda'_0/\lambda_0)_A$. Thus the Doppler

† Similarly, $T_B-\phi(T_B) = t_1 = p_{12}f(T_B) = p_{12}(t'_B)$

shifts observed at the same local times are identical—a particular case of the equivalence of A and B

22. As an example, suppose that A observes the Doppler shift of the light from B to be *constant in time*, $s(t_2) = \text{const} = s$ Then

$$dt'_B = \frac{dt_2}{s},$$

or, since t'_B and t_2 vanish together,

$$t'_B = \frac{t_2}{s}$$

Hence $$s = p'_{21}(t'_B) = p'_{21}(t_2/s),$$

or, integrating, $$p_{21}(t_2/s) = t_2$$

Hence $$p_{21}(\xi) = s\xi, \tag{14}$$

and since p_{12} and p_{21} are inverse functions,

$$p_{12}(\xi) = \xi/s.$$

Hence $$f^{-1}(\xi) = \tfrac{1}{2}\{p_{12}(\xi)+p_{21}(\xi)\} = \tfrac{1}{2}\xi\left(s+\frac{1}{s}\right) = \frac{s^2+1}{2s}\xi,$$

or $$f(\xi) = \frac{2s}{s^2+1}\xi \tag{15}$$

Hence $$\begin{aligned}\phi(\xi) &= \tfrac{1}{2}[p_{21}f(\xi)-p_{12}f(\xi)] \\ &= \frac{1}{2}\left[s\frac{2s}{s^2+1}-\frac{1}{s}\frac{2s}{s^2+1}\right]\xi \\ &= \frac{s^2-1}{s^2+1}\xi \end{aligned} \tag{16}$$

This determines the time-distance function ϕ Accordingly, in terms of the coordinates R_B, T_B used by A, A's observations of the time-distance behaviour of B are given by

$$R_B = c\frac{s^2-1}{s^2+1}T_B$$

Hence in terms of the coordinates used by himself, A judges B's distance to be increasing at the uniform rate

$$\frac{dR_B}{dT_B} = c\frac{s^2-1}{s^2+1}$$

Radial velocity

23. If A's observation of B, expressed in terms of coordinates, are stated in the form $R_B = c\phi(T_B)$, A defines the 'radial velocity' of B as dR_B/dT_B, and denotes this by $V(T_B)$ In the example just considered, in which s was constant in time, $V(T_B)$ is constant and is given by

$$V = c\frac{s^2-1}{s^2+1}, \tag{17}$$

or

$$s = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}} \tag{18}$$

This is a well-known formula in the special theory of relativity Here we have deduced it merely from the equivalence of A and B, that is to say the identity of A's description of his observations of B with B's description of his observations of A We have thus proved the fundamental result that if two particle-observers are equivalent, and if the Doppler effects they observe on one another are constant in time, then they have a uniform relative radial velocity, or are in uniform relative radial motion, in terms of the coordinates we have defined In establishing this result we have supposed that A can recognize an atom on B as being identical with a certain species of atom at himself, so as to be able to compare wave-lengths † But we have not assumed the concept of the transport of rigid length-scales, or introduced any postulate as to the 'velocity of light', which we have not even defined

24 It follows that to the extent to which a nebula exhibits a red-shift constant in time, if that nebula is 'equivalent' to ourselves, then it is moving relatively to us in terms of the coordinates commonly used in physics, and moreover at a constant relative velocity More generally, it can be shown that if $s \neq 1$, whether s is constant in time or not, then V as defined is non-zero Any suggested explanation of the red-shift observed in a distant nebula as 'due to' other than relative velocity can only be maintained by abandoning the 'equivalence' of the distant nebula to ourselves, in the sense in which we have defined 'equivalence'

† This recognition is only possible in virtue of the discreteness of atomic properties If the world was composed of atoms shading continuously into one another, no such recognition would be possible

25. In the case of a constant Doppler shift s, we have now by (16), (15), (14), and (18)

$$\phi(\xi) = V\xi, \tag{19}$$

$$f(\xi) = (1-V^2/c^2)^{\frac{1}{2}}\xi, \tag{20}$$

$$p_{12}(\xi) = \left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}\xi, \qquad p_{21}(\xi) = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}}\xi \tag{21}$$

Relation (20), written in the form $t'_B = f(T_B)$, gives

$$t'_B = (1-V^2/c^2)^{\frac{1}{2}}T_B \tag{22}$$

and so fixes the clock-behaviour at B in terms of A's coordinate time T_B for the same event. In the world-wide instant $t = T_B$ in A's coordinates, B's clock reads the smaller time $(1-V^2/c^2)^{\frac{1}{2}}T_B$. Thus judged by A's clock, B's clock is running slow. This is a consequence merely of the equivalence of the two particle-observers as regards their observations of one another.

We proceed now with the general theory, not restricting ourselves to constant Doppler shifts or to uniform relative motion.

The velocity of light

26. Consider again the signals scheme of Fig. 3. We have

$$\tfrac{1}{2}c(t_2-t_1) = R_B = c\phi(T_B),$$

$$\tfrac{1}{2}(t_2+t_1) = T_B,$$

whence

$$T_B-\phi(T_B) = t_1 \tag{23}$$

Write this in the form

$$\frac{c\phi(T_B)}{T_B-t_1} = c$$

or

$$\frac{R_B}{T_B-t_1} = c \tag{24}$$

By A's clock, T_B-t_1 is the interval occupied by the signal in overtaking B, since T_B is the epoch, by A's clock, of the event E_B which was the reception by B of the signal which left A at epoch t_1 by A's clock; further, R_B is the distance, by A's clock, traversed by this signal. We define the ratio

$$\frac{\text{distance described by the signal}}{\text{interval occupied by the signal}}$$

as the mean outward signal-velocity between the event t_1 at A and the event E_B at B. By (24) this is constant and equal to c. Similarly, we have

$$\frac{R_B}{t_2-T_B} = c, \tag{25}$$

whence the mean inward signal velocity between E_B and t_2 is also equal to c

Now consider Fig 4 We have

$$\tfrac{1}{2}c(t'_4-t'_3) = R'_A = c\phi(T'_A),$$

$$\tfrac{1}{2}(t'_4+t'_3) = T'_A,$$

whence
$$\frac{R'_A}{T'_A-t'_3} = c = \frac{R'_A}{t'_4-T'_A}. \tag{26}$$

Thus B's calculation of the outward and inward mean signal-velocities, similarly defined, also comes to c The numbers assigned by A or by B, to inward or outward mean signal velocities, are thus equal and independent of the events E_B and E_A concerned The number c was an arbitrary number conventionally adopted, by agreement, by A and by B as a means of constructing coordinates (epochs and distances of distant events) out of observations We now see that if they define velocities as quotients of conventional distance-measures by conventional epoch-differences, they will attach constant and equal numbers to the signal-velocity This velocity they will call the velocity of light

We see that the famous 'postulate of the constancy of the velocity of light' is at bottom a convention If two particle-observers are equivalent in the sense that they adopt similar rules for converting observations into conventional coordinates, and if their observations on one another are then described in the same terms, they are necessarily led to attach equal numbers to their reckoning of signal-velocities

The recognition of this state of affairs is due to the genius of Einstein Einstein's procedure was, however, to assume that a meaning could be attached to the transport of identical clocks and rigid length-scales, and then to assume as a physical fact that the resulting values of the velocity of light, in either direction, would be equal. This procedure does not, however, devise tests for verifying observationally that the clocks and scales carried are identical Moreover he assumed *uniform* relative velocity We have stated an experimental procedure for verifying the equivalence of particle-observers in terms of observations that could be actually carried out We have by this means attached numerical measures to lengths or distances obtained by clock-readings only. We have not assumed uniform relative velocity Comparing our procedure with Einstein's, we may

say that we have given a definition of equality of distances assigned by observers in relative motion, and so have defined rigid length-scales at a distance (We shall see shortly that our definition of length or distance coincides with that used in ordinary physics) We have then found that the 'constancy of the velocity of light' is an *a posteriori* consequence of the equivalence of the observers

Generalized Lorentz formulae in one dimension

27. So far we have not required to consider whether A and B were 'in relative motion along a straight line', or to attach a meaning to this phrase The number V, namely $c\phi'(T_B)$, is what we ordinarily call the radial component of velocity of B relative to A We shall now restrict attention to the case in which a meaning can be attached to saying that B is in a 'constant direction' relative to A, and we proceed to consider events E in line with B and A. We propose to relate the coordinates assigned to such events by A and by B

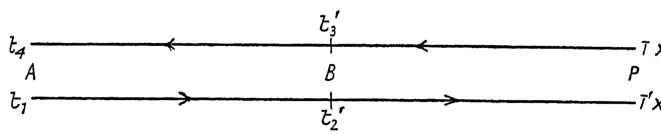


FIG 5 Derivation of generalized Lorentz formulae for one dimension

Consider the scheme of signals depicted in Fig 5 A sends at time t_1 by his clock a signal which passes over B at time t_2' by B's clock, is reflected at a distant particle P in line with A and B, and passes over B at time t_3' by B's clock, returning to A at time t_4 by A's clock The epochs t_1, t_4 are directly observed by A, the epochs t_2', t_3' by B A assigns to the event E of reflection at P the epoch T and distance X, where

$$T = \tfrac{1}{2}(t_4+t_1), \qquad X = \tfrac{1}{2}c(t_4-t_1), \tag{27}$$

whilst B assigns to the same event E the epoch T' and distance X', where

$$T' = \tfrac{1}{2}(t_3'+t_2'), \qquad X' = \tfrac{1}{2}c(t_3'-t_2') \tag{28}$$

Hence

$$T+X/c = t_4, \qquad T-X/c = t_1 \tag{29}$$

$$T'+X'/c = t_3', \qquad T'-X'/c = t_2' \tag{30}$$

We have now expressed certain functions of T, X, T', X' in terms of observations t_1, t_4, t_2', t_3' of events at A and B only These can now be connected from our knowledge of the relative motion of A and B

For A's observations on B are consistent with the statement that if t_2 is A's epoch for t_2' at B, a signal leaving A at t_1 and reaching B at

t_2, in A's measures, describes a distance $c(t_2-t_1)$, which is equal to B's distance $c\phi(t_2)$ (in A's reckoning) at the event t_2'. Hence

$$t_2-t_1 = \phi(t_2)$$

(This is a particular case of equation (23).) Similarly, B writes down

$$t_4'-t_3' = \phi(t_4'),$$

where t_4' is the epoch assigned by B to the event t_4 at A. Hence

$$t_2-\phi(t_2) = t_1,$$

or by (7′)

$$t_1 = p_{12}f(t_2)$$

$$= p_{12}(t_2'), \tag{31}$$

since t_2' is an event at B. Similarly,

$$t_4' - \phi(t_4') = t_3',$$

or by (7′)

$$t_3' = p_{12}f(t_4')$$

$$= p_{12}(t_4), \tag{32}$$

since t_4 is an event at A.

Hence from (29) and (30), introduced in (31) and (32),

$$T-X/c = p_{12}(T'-X'/c),$$
$$T'+X'/c = p_{12}(T+X/c).$$

These may be rewritten in the more instructive forms

or

$$\left.\begin{aligned} T'+X'/c &= p_{12}(T+X/c), \\ T'-X'/c &= p_{21}(T-X/c), \\ T+X/c &= p_{21}(T'+X'/c), \\ T-X/c &= p_{12}(T'-X'/c). \end{aligned}\right\} \tag{33}$$

These determine B's assignments T', X' in terms of A's assignments T, X and conversely.

28. In the case in which the Doppler effect of B observed at A is constant, its value s defines a velocity V equal to dR_B/dT_B or dR_A'/dT_A', the functions p_{12}, p_{21} are given by (21), and we have accordingly, by (33),

$$T'+X'/c = \left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}(T+X/c), \qquad T'-X'/c = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}}(T-X/c).$$

These yield

$$T'^2-X'^2/c^2 = T^2-X^2/c^2, \tag{34}$$

and moreover on solution for T, X or T', X' give

$$T' = \frac{T-VX/c^2}{(1-V^2/c^2)^{\frac{1}{2}}}, \qquad X' = \frac{X-VT}{(1-V^2/c^2)^{\frac{1}{2}}}, \tag{35}$$

$$T = \frac{T'+VX'/c^2}{(1-V^2/c^2)^{\frac{1}{2}}}, \qquad X = \frac{X'+VT}{(1-V^2/c^2)^{\frac{1}{2}}} \tag{36}$$

These are the well-known Lorentz formulae in one dimension We have made their derivation turn on the determination of the behaviour of B's clock in A's experience and of A's clock in B's experience

29. In the general case we have

$$T'^2 - X'^2/c^2 = p_{12}(T+X/c)p_{21}(T-X/c),$$
$$T^2 - X^2/c^2 = p_{21}(T'+X'/c)p_{12}(T'-X'/c)$$

In general T^2-X^2/c^2 and $T'^2-X'^2/c^2$ are unequal

We have supposed the event E to be outside A and B on the same side of A as B It is readily verified that with a simple convention as to the sign of X (positive in the direction AB, negative in the direction BA), the formulae hold generally The reader will find it an instructive exercise to derive them when E lies between A and B, in that case A and B must send oppositely directed signals which arrive simultaneously at E

30. For two neighbouring events (X, T), $(X+dX, T+dT)$ or (X', T'), $(X'+dX', T'+dT')$, we have

$$\left.\begin{aligned} dT'+dX'/c &= p'_{12}(T+X/c)(dT+dX/c), \\ dT'-dX'/c &= p'_{21}(T-X/c)(dT-dX/c), \\ dT+dX/c &= p'_{21}(T'+X'/c)(dT'+dX'/c), \\ dT-dX/c &= p'_{12}(T'-X'/c)(dT'-dX'/c) \end{aligned}\right\} \tag{37}$$

Multiplying together in pairs, we have from the first and fourth,

$$p'_{12}(T'-X'/c)(dT'^2-dX'^2/c^2) = p'_{12}(T+X/c)(dT^2-dX^2/c^2), \tag{38}$$

and from the second and third

$$p'_{21}(T'+X'/c)(dT'^2-dX'^2/c^2) = p'_{21}(T-X/c)(dT^2-dX^2/c^2)$$

These state the same fact, since

$$\left.\begin{aligned} \frac{dT'+dX'/c}{dT+dX/c} &= p'_{12}(T+X/c) = \frac{1}{p'_{21}(T'+X'/c)}, \\ \frac{dT'-dX'/c}{dT-dX/c} &= p'_{21}(T-X/c) = \frac{1}{p'_{12}(T'-X'/c)} \end{aligned}\right\} \tag{39}$$

Thus dT^2-dX^2/c^2 is a covariant The relations show that $dT'^2-dX'^2/c^2$ vanishes when dT^2-dX^2/c^2 vanishes Thus if the particle P is moving with the velocity $dX/dT = c$ in the experience of A, it is moving with the same velocity $dX'/dT' = c$ in the experience of B

We have, moreover, that the quadratic form

$$[p'_{12}(T+X/c)p'_{21}(T-X/c)]^{\frac{1}{2}}(dT^2-dX^2/c^2)$$

is numerically equal to the quadratic form

$$[p'_{12}(T'-X'/c)p'_{21}(T'+X'/c)]^{\frac{1}{2}}(dT'^2-dX'^2/c^2).$$

Bearing in mind that A and B are using (relatively) different sign conventions for their X-measures (the one, X positive when towards the other observer, the other, X positive when away from the first observer) we see that the quadratic form in question is an invariant, taking the same value in A's assignments and B's When A and B are in uniform relative velocity, p'_{12} and p'_{21} are reciprocal constants, and the quadratic form reduces to dT^2-dX^2/c^2 It is readily shown that the form dT^2-dX^2/c^2 is equal to the form $dT'^2-dX'^2/c^2$ only when A and B are in uniform relative velocity

'*Cosmic*' *time*

31. Since the formulae we have derived relating to coordinates used by two observers coincide, for uniform relative velocity, with the Lorentz formulae, it follows that the velocity V occurring in the Lorentz formulae as specifying the relative velocity of the two observers is equal to what we have called dX_B/dT_B This is accordingly what we mean by 'velocity' in ordinary physics, as for example in the electromagnetic theory or in Einstein's addition formulae (see below) It is a differential coefficient constructed by A out of his own measures, namely the constructs X_B and T_B It should be carefully distinguished from dX_B/dt'_B, which may be called A's reckoning of the 'cosmic' velocity of B, using B's measure of time, the 'cosmic' time t'_B of events at B The latter is a 'mixed' differential coefficient, employing partly A's measures, partly B's It can be shown that as $dX_B/dT_B \to c$, $dX_B/dt'_B \to \infty$, thus if a particle is moving with the velocity of light, in the experience of an observer A, its 'cosmic' velocity is infinite

For consider formulae (39) applied to events at B For such events, $T' = t'_B$, $X'_B = 0$, $T = T_B$, $X = X_B$, and we have accordingly

$$\frac{dt'_B}{dT_B+dX_B/c} = \frac{1}{p'_{21}(t'_B)}, \qquad \frac{dt'_B}{dT_B-dX_B/c} = \frac{1}{p'_{12}(t'_B)}$$

Now let $p'_{21}(t'_B) \to \infty$ From earlier results, this means that the Doppler effect observed by A on B at epoch $p_{21}(t'_B)$ by A's clock tends to infinity, $s \to \infty$ Then $dt'_B \to 0$, and since $p'_{12}(t'_B)$ will be in general finite, the second equation shows that $dX_B/dT_B \to c$ † It follows that $dX_B/dt'_B \to \infty$ Thus an infinite Doppler shift at B observed by A implies that B is moving with the velocity assigned to light, in A's experience, whilst its 'cosmic' velocity is infinite Other coordinates may be differentiated with regard to 'cosmic' time, t'_B, yielding other varieties of 'cosmic' velocities, but in all cases these tend to infinity as the velocity of B, in the experience of some other equivalent observer, approaches that of light Such 'cosmic' velocities are frequently employed in relativistic cosmology, and have occasionally been interpreted as meaning that particles are moving with speeds exceeding that of light, in cases where an analogue of dX_B/dt'_B is very large This is a misunderstanding Any finite value of dX_B/dt'_B, however large, corresponds to a coordinate velocity dX_B/dT_B numerically less than c The meaning of $dt'_B \to 0$ as $dX_B/dT_B \to c$ is that when B appears to A to be moving with the speed of light, B's clock stands still in A's view

The coordinate time T assigned to an event by A, and the velocity dX/dT, are the time and velocity commonly used in physics and in ordinary life For example this T is the time used in timing a race, stating athletic records, or arranging a railway time-table, the world-wide instant $t = T$ is what we ordinarily mean by simultaneity Every individual observer can, and in fact does, adopt *for his own purposes* the Newtonian concept of a world-wide time Whether this is 'evenly flowing' or not is irrelevant and meaningless It suffices that the observer can unambiguously attach an epoch T to every event he can observe, and that this epoch T has a definite significance in terms of light-signals, events with the same T are what we ordinarily call simultaneous It is only when he compares his experiences with those of a second observer that he may find, and in general will find, that this other observer will attach a variety of epoch-numbers T' to events labelled T by the first observer To the first observer T is always the epoch of occurrence of the event as reckoned by the observer's own clock

† This can also be established by starting with $\phi'(T_B) = 1$, using

$$p_{12}(t'_B) = p_{12} f(T_B) = T_B - \phi(T_B),$$

and differentiating

Alternative deduction of the Lorentz formulae for uniform motion in one dimension

32. The following method may be appreciated by the reader who finds our analysis in terms of p_{12} functions a little indirect It is a shorter method of deriving the Lorentz formulae for *uniform* motion in one dimension

Equations (5) and (6) connecting two observers A and B are derived exactly as before We now assume that $c\phi(\xi)$ is the linear function $V\xi$, $(0 < V < c)$, and have accordingly

$$f(y) = x(1-V/c),$$
$$f(x) = y(1+V/c)$$

Hence
$$f\left(\frac{f(x)}{1+V/c}\right) = x(1-V/c)$$

Put
$$\frac{f(x)}{1+V/c} = \psi(x) \tag{40}$$

Then
$$\psi(\psi(x)) = \frac{1-V/c}{1+V/c}x$$
$$= \alpha^2 x, \tag{41}$$

say, where
$$\alpha^2 = \frac{1-V/c}{1+V/c} \qquad (0 < \alpha^2 < 1). \tag{42}$$

Operate with ψ on both sides of (41) Then

$$\alpha^2\psi(x) = \psi(\alpha^2 x)$$

A particular solution is $\psi(x) = \alpha x$ Write in general†

$$\psi(x) = \alpha x\Psi(x)$$

Then
$$\alpha^3 x\Psi(x) = \alpha^3 x\Psi(\alpha^2 x),$$

or
$$\Psi(x) = \Psi(\alpha^2 x)$$

Hence
$$\Psi(x) = \Psi(\alpha^{2n}x) \qquad (n \text{ integral})$$

Let $n \to \infty$ Then since $0 < \alpha < 1$, $\alpha^{2n} \to 0$ and

$$\Psi(x) = \Psi(0),$$

assuming Ψ to be continuous Hence $\psi(x) = \alpha x\Psi(0)$ Introducing in (41) we have $\Psi(0) = \pm 1$ Clearly the upper sign is required, and so $\psi(x) = \alpha x$ Hence

$$f(x) = (1+V/c)\left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}x$$
$$= (1-V^2/c^2)^{\frac{1}{2}}x \tag{43}$$

† This device, amongst the many possible procedures, was suggested to me by Mr G J Whitrow

Accordingly $t'_B = (1-V^2/c^2)^{\frac{1}{2}}T_B, \qquad t_A = (1-V^2/c^2)^{\frac{1}{2}}T'_A$ (43')

Relations (43) enable A and B to calculate the epochs of events at A or B as directly *recorded* by A or B in terms of the epochs assigned to the same event by B or A, respectively They are essentially more primitive than the Lorentz formulae, which we proceed to deduce from them

33. We first notice a result of interest The relation

$$T_B - R_B/c = t_A,$$

already deduced, is in this case

$$T_B - t_A = R_B/c,$$

and so is compatible with the statement that the light covers the distance R_B with speed c in a time $T_B - t_A$ It may also be written, since $R_B = VT_B$,

$$T_B - t_A = \frac{Vt_A}{c-V},$$

which is compatible with the statement that the light makes up its starting handicap Vt_A (the distance of B at epoch t_A) in time $T_B - t_A$, at the rate $c-V$ Thus $c-V$ is A's reckoning of the speed of the signal relative to B This statement must not be confused with the statement that relative to B, *in B's reckoning*, the speed of the signal is c

34. We now derive the one-dimensional Lorentz formulae, employing again the notation of Fig 5 We repeat the meaning of the symbols for convenience A signal leaves A at t_1 by A's clock, passes over B at t'_2 by B's clock, is reflected at a distant particle P, passes B on the return journey at time t'_3 by B's clock, and returns to A at t_4 by A's clock The distant particle P is in line with A and B The numbers t_1, t'_2, t'_3, t_4 are actual observations From these observations A and B assign epochs and distances to the event E of reflection at P

A writes down $\qquad T = \tfrac{1}{2}(t_4+t_1), \qquad X = \tfrac{1}{2}c(t_4-t_1),$

B writes down

$$T' = \tfrac{1}{2}(t'_3+t'_2), \qquad X' = \tfrac{1}{2}c(t'_3-t'_2)$$

Hence

$$T+X/c = t_4, \qquad T-X/c = t_1,$$

$$T'+X'/c = t'_3, \qquad T'-X'/c = t'_2$$

By the formulae (43) already found, A can compute by his clock the epochs t_2 and t_3 of the events t'_2 and t'_3 at B He finds

$$t'_2 = (1-V^2/c^2)^{\frac{1}{2}}t_2, \qquad t'_3 = (1-V^2/c^2)^{\frac{1}{2}}t_3$$

But since his observations have already found B to be in motion with uniform velocity $dR_B/dT_B = V$, he can write down

$$c(t_2-t_1) = Vt_2, \qquad c(t_4-t_3) = Vt_3,$$

or

$$t_2 = \frac{ct_1}{c-V}, \qquad t_3 = \frac{ct_4}{c+V}$$

Combining these relations,

$$T'+X'/c = t_3' = (1-V^2/c^2)^{\frac{1}{2}}t_3$$
$$= (1-V^2/c^2)^{\frac{1}{2}}\frac{ct_4}{c+V} = \left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}(T+X/c),$$
$$T'-X'/c = t_2' = (1-V^2/c^2)^{\frac{1}{2}}t_2$$
$$= (1-V^2/c^2)^{\frac{1}{2}}\frac{ct_1}{c-V} = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}}(T-X/c)$$

Multiplying together we have

$$T'^2-X'^2/c^2 = T^2-X^2/c^2, \tag{44}$$

and solving we have

$$X' = \frac{X-VT}{(1-V^2/c^2)^{\frac{1}{2}}}, \qquad T' = \frac{T-VX/c^2}{(1-V^2/c^2)^{\frac{1}{2}}}, \tag{45}$$

$$X = \frac{X'+VT'}{(1-V^2/c^2)^{\frac{1}{2}}}, \qquad T = \frac{T'+VX'/c^2}{(1-V^2/c^2)^{\frac{1}{2}}} \tag{45'}$$

Lorentz-Fitzgerald contraction

35 If we take two sequences of events $(T', X_1'), (T', X_2')$ in B's experience, where $X_2'-X_1' = \text{const} = l$ and T' is variable, A observes these sequences as $(T_1, X_1), (T_2, X_2)$ given in terms of T', X_1', X_2' by the above formulae For the given common epoch T' by B's clock, A finds

$$X_2-X_1 = \frac{X_2'-X_1'}{(1-V^2/c^2)^{\frac{1}{2}}} = \frac{l}{(1-V^2/c^2)^{\frac{1}{2}}},$$

and thus assigns a greater value to the distance-separation of the events But if A picks out of the sequences events having a common epoch T by A's clock, A finds then

$$X_2'-X_1' = \frac{X_2-X_1}{(1-V^2/c^2)^{\frac{1}{2}}}$$

or

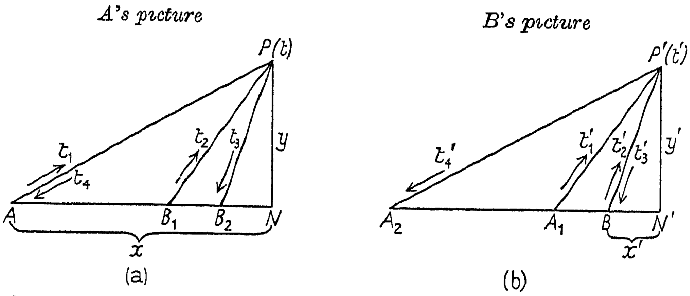
$$X_2-X_1 = l(1-V^2/c^2)^{\frac{1}{2}}$$

The two sequences of events in B's experience may be taken to define a rigid body or rod of invariable length l moving with B in the direction of B's motion At epochs which A considers simultaneous, of epoch T, the length A assigns to the rod is $l(1-V^2/c^2)^{\frac{1}{2}}$ Thus in A's

experience the rod appears to have a shorter length than in B's experience, when each evaluates its 'length' or the positions of its ends at epochs simultaneous for himself † This phenomenon is called the 'Lorentz-Fitzgerald contraction'

Lorentz formulae in three dimensions

36 To generalize the Lorentz formulae so as to connect A's and B's observations of an event E at a particle P not in line with A and B we may proceed as follows The analysis is restricted to the case where A and B are in uniform relative motion



FIGS 6a, 6b Derivation of Lorentz formulae in three dimensions

Let A send a signal at time t_1 by A's clock Let this be reflected at P and return to A at time t_4 by A's clock The returning signal passes over B at time t_3' by B's clock, and without difficulty we may suppose that B dispatched at time t_2' by B's clock a signal which arrived at P simultaneously with A's original signal A and B then construct the diagrams shown in Figs 6a and 6b

Further, A *assumes* tentatively that he may make his calculations as if the signal from and to B is propagated rectilinearly along what to him (A) are the 'cross-country' paths B_1P, PB_2, and B makes a similar assumption about what to him are the cross-country paths A_1P', $P'A_2$ If these assumptions are not justified, they will show themselves as inconsistencies in the ensuing analysis, as we shall see A and B now assume Euclidian geometry for the figures they have constructed, in particular the Pythagorean theorem This allows them to assign numbers y, y' to the 'lengths' of the 'perpendiculars' PN, $P'N'$ in their diagrams from P (or P') to the line of motion AB, and numbers x, x' to the projections of AP, BP' on this line All the 'lengths' AP, PB_1, PB_2, BP', $P'A_1$, $P'A_2$ are known from the light-

† See Supplementary Notes, p 59

times, as are AB_1, AB_2, BA_1, BA_2, and so A and B write down the following systems of equations †

A's system	*B's system*
$x^2+y^2 = c^2(t-t_1)^2$	$x'^2+y'^2 = c^2(t'-t_2')^2$
$x^2+y^2 = c^2(t_4-t)^2$	$x'^2+y'^2 = c^2(t_3'-t')^2$
$(x-Vt_2)^2+y^2 = c^2(t-t_2)^2$	$(x'+Vt_1')^2+y'^2 = c^2(t'-t_1')^2$
$(x-Vt_3)^2+y^2 = c^2(t_3-t)^2$	$(x'+Vt_4')^2+y'^2 = c^2(t_4'-t')^2$ (46)

In these equations, t_1, t_4, t_2', t_3' are observed numbers, whilst t_1', t_4', t_2, t_3 are the corresponding epochs (of events at A and B respectively) assigned by B and A. They are known from the relations already established,

$$t_2' = t_2(1-V^2/c^2)^{\frac{1}{2}}, \qquad t_3' = t_3(1-V^2/c^2)^{\frac{1}{2}},$$

$$t_1 = t_1'(1-V^2/c^2)^{\frac{1}{2}}, \qquad t_4 = t_4'(1-V^2/c^2)^{\frac{1}{2}} \tag{47}$$

Equations (46) thus provide 8 relations for the 6 unknowns (t, x, y), (t', x', y') in terms of the 4 observed numbers t_1, t_4, t_2', t_3'. If we eliminate t_1, t_4, t_2', t_3', we shall be left with 4 relations connecting t, x, y, t', x', y'.

From the third of A's equations (46) we have

$$\frac{y^2}{c^2} = \left[\left(t-\frac{x}{c}\right)-t_2\left(1-\frac{V}{c}\right)\right]\left[t+\frac{x}{c}-t_2\left(1+\frac{V}{c}\right)\right]$$

Substituting for t_2 in terms of t_2' by means of the first of equations (47), and then substituting for t_2' from the first of B's equations (46), we have

$$\frac{y^2}{c^2} = \left[\left(t-\frac{x}{c}\right)-\alpha\left(t'-\frac{(x'^2+y'^2)^{\frac{1}{2}}}{c}\right)\right]\left[t+\frac{x}{c}-\alpha^{-1}\left(t'-\frac{(x'^2+y'^2)^{\frac{1}{2}}}{c}\right)\right], \tag{48}$$

where

$$\alpha = \left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}$$

Similarly, the last of A's equations (46), the second of (47), and the second of B's equations (46) give

$$\frac{y^2}{c^2} = \left[\left(t-\frac{x}{c}\right)-\alpha\left(t'+\frac{(x'^2+y'^2)^{\frac{1}{2}}}{c}\right)\right]\left[t+\frac{x}{c}-\alpha^{-1}\left(t'+\frac{(x'^2+y'^2)^{\frac{1}{2}}}{c}\right)\right] \tag{48'}$$

Subtracting the last two equations we find easily

$$t' = \tfrac{1}{2}\left[\alpha\left(t+\frac{x}{c}\right)+\alpha^{-1}\left(t-\frac{x}{c}\right)\right]$$

† We are here temporarily using small letters for the calculated coordinates of distant events. See also Supplementary Notes, p. 59

Similarly, we can prove that

$$t = \tfrac{1}{2}\left[\alpha^{-1}\left(t'+\frac{x'}{c}\right)+\alpha\left(t'-\frac{x'}{c}\right)\right]$$

These give $$t'+\frac{x'}{c} = \alpha\left(t+\frac{x}{c}\right), \qquad t'-\frac{x'}{c} = \alpha^{-1}\left(t-\frac{x}{c}\right), \tag{49}$$

which yield relations (45), (45′), which are now seen to hold good for point-events *off the line of motion* Further, (48) yields, on using (49),

$$\frac{y^2}{c^2} = \left[\left(t'-\frac{x'}{c}\right)-\left(t'-\frac{(x'^2+y'^2)^{\frac{1}{2}}}{c}\right)\right]\left[\left(t'+\frac{x'}{c}\right)-\left(t'-\frac{(x'^2+y'^2)^{\frac{1}{2}}}{c}\right)\right]$$
$$= y'^2/c^2, \tag{50}$$

so that $y = y'$ There are 4 equations of the type (48), obtained by eliminating the 4 variables t_1, t_4, t_2', t_3' from the 8 equations (46), providing 4 relations between t, x, y and t', x', y', of which we have found 3 The other relation reduces as above to $y = y'$ This establishes the Lorentz formulae for events in two dimensions outside the line of relative motion, and the extension to three dimensions requires no comment It is worthy of note that equations (49) are identical with the one-dimensional Lorentz formulae previously obtained, but this could by no means have been taken for granted †

37 We have now completed our derivation of the Lorentz formulae Each observer using coordinates (t, x, y, z), where t is the epoch of the event by his own clock, can combine spatial coordinates x, y, z by the rules of Euclidian geometry, and can make diagrams or maps valid for any world-wide instant of Newtonian time t by his clock The Lorentz formulae are simply means of passing from one Euclidian space and Newtonian time to those of a second equivalent observer in uniform relative motion

38. The relationship of A's and B's spaces is often described by constructing a four-dimensional manifold 'space-time' in which an event, represented by a point of the manifold, is described as of coordinates (t, x, y, z) with reference to one set of axes, (t', x', y', z') with respect to another This way of regarding the matter has been historically of great importance, but it fails to emphasize the essential point of the unidirectional 'flow of time' for each individual observer, and it involves a spurious identification of time and space, which

† This, however, is often done in text-books

require later to be once more separated when formulae come to be physically interpreted In our development we shall find it more convenient and more insight-giving to avoid the use of space-time and to treat space-coordinates and time-coordinates on the different footings on which they first appear

Linear systems of equivalent observers in non-uniform relative motion

39. In preceding sections we have found (*a*) *generalized* one-dimensional Lorentz formulae for the coordinates (T, X) (T', X') of any event E in line with the equivalent observers A and B, (*b*) three-dimensional Lorentz formulae for *any* event E in the case of uniform relative motion of A and B, $X_B = VT_B$, $X'_A = -VT'_A$ It is natural to try to find the generalized Lorentz formulae in three dimensions, for any event E not necessarily lying in the line AB, in the general case $X_B = c\phi(T_B)$, $X'_A = -c\phi(T'_A)$ If we attempt to do so, using Euclidian geometry as in Figs 6 a, 6 b, and assuming rectilinear 'cross-country' light-paths† PB_1, PB_2, etc , we are easily led to 4 equations of the type (48) which should define t', x', y' in terms of t, x, y Actually any one of the 4 equations is redundant, and proves to be inconsistent with the other 3 This shows that a false hypothesis has been made, and if, as we suppose, the geometry may be chosen arbitrarily, the only hypothesis in question is the rectilinearity of the cross-country light-paths These must therefore be curved In other words, in the presence of two particle observers A and B undergoing relative acceleration, A must assume curvatures of light-paths save perhaps for paths terminating or starting from himself Following Einstein, we may suppose that to say that A and B are undergoing relative acceleration is equivalent to saying that they are moving under the influence of, or in the presence of, a gravitational field We infer, therefore, that a gravitational field affects the curvature of light-paths, which is one of Einstein's fundamental results here obtained by a different method

It should be possible to determine the curvature of these light-paths from the condition of equivalence of the two observers, and hence to obtain the generalized Lorentz formulae for any event P for the case of general accelerated motion $X_B = c\phi(T_B)$, $X'_A = -c\phi(T'_A)$ I have not succeeded in solving this problem Instead of discussing possible ways of solving it (which may be left for future research),

† Described with speed c

I shall now proceed to discuss in detail one-dimensional sets of equivalent observers, by means of the one-dimensional Lorentz formulae already found These will be found to give insight, in a preliminary way, into the cosmological problem

40. Let three observers, A, B, C, moving in the same straight line and having at one time been coincident, be such that $A \equiv B$ and $A \equiv C$ We want to find the necessary and sufficient conditions that $B \equiv C$

41. First, to find the necessary conditions, let us suppose $B \equiv C$ and determine the consequences Then for the three pairs (A, B), (B, C), (A, C) there exist pairs of inverse functions (p_{12}, p_{21}), (p_{23}, p_{32}), (p_{13}, p_{31}) such that if (T, X), (T', X'), (T'', X'') are the three descriptions of an event E in line with ABC by A, B, and C respectively, then

$$T'+X'/c = p_{12}(T+X/c), \qquad T'-X'/c = p_{21}(T-X/c),$$
$$T''+X''/c = p_{23}(T'+X'/c), \qquad T''-X''/c = p_{32}(T'-X'/c),$$
$$T''+X''/c = p_{13}(T+X/c), \qquad T''-X''/c = p_{31}(T-X/c)$$

From these we have the operational identities

$$p_{13} = p_{23}\,p_{12}, \tag{51}$$
$$p_{31} = p_{32}\,p_{21} \tag{52}$$

Taking the inverse of the last identity† we have

$$p_{13} = p_{12}\,p_{23}$$

Hence

$$p_{12}\,p_{23} = p_{23}\,p_{12} \tag{53}$$

The operators p_{12} and p_{23} must therefore commute When they commute, their 'product' determines p_{13} Similarly we prove

$$p_{23} = p_{13}\,p_{21} = p_{21}\,p_{13} \tag{54}$$

and

$$p_{12} = p_{13}\,p_{32} = p_{32}\,p_{13}, \tag{55}$$

which are also consequences of the condition $p_{13} = p_{12}\,p_{23} = p_{23}\,p_{12}$ These are the necessary conditions

42. We now prove that these conditions are sufficient If $A \equiv B$, their observations are related by

$$X_B = c\phi_{12}(T_B), \qquad T'_B = f_{12}(T_B),$$
$$X'_A = +c\phi_{21}(T'_A), \qquad T_A = f_{21}(T'_A),$$

where

$$\phi_{21} \equiv -\phi_{12}, \qquad f_{12} \equiv f_{21}$$

† If f_1, f_2, , f_n are any operators possessing inverses and $F = f_1 f_2 \;\; f_n$, then $F^{-1} = f_n^{-1} \;\; f_2^{-1} f_1^{-1}$ This is easily proved.

Further, inverse functions p_{12}, p_{21} exist such that for any ζ

$$\zeta - \phi_{12}(\zeta) = p_{12} f_{12}(\zeta),$$
$$\zeta - \phi_{21}(\zeta) = p_{21} f_{12}(\zeta)$$

These follow from § 20, after a slight change of notation

Again, if $A \equiv C$, their observations are related by

$$X_C = c\phi_{13}(T_C), \qquad T''_C = f_{13}(T_C),$$
$$X''_A = c\phi_{31}(T''_A), \qquad T_A = f_{31}(T''_A),$$

where $$\phi_{31} \equiv -\phi_{13}, \qquad f_{13} \equiv f_{31},$$

and inverse functions p_{13}, p_{31} exist such that

$$\zeta - \phi_{13}(\zeta) = p_{13} f_{13}(\zeta),$$
$$\zeta - \phi_{31}(\zeta) = p_{31} f_{13}(\zeta)$$

Let B observe C to have the motion

$$X'_C = c\phi_{23}(T'_C),$$

and let C observe B to have the motion

$$X''_B = c\phi_{32}(T''_B)$$

Further, let B observe C's clock to have the behaviour

$$T''_C = f_{23}(T'_C),$$

and let C observe B's clock to have the behaviour

$$T'_B = f_{32}(T''_B)$$

We require to prove that if $p_{12}p_{31} = p_{31}p_{12}$, then

$$\phi_{23} \equiv -\phi_{32}, \qquad f_{23} = f_{32}$$

Consider events at C For these, $X''_C = 0$ Hence we have the relations

$$T_C + X_C/c = p_{21}(T'_C + X'_C/c) = p_{31}(T''_C),$$
$$T_C - X_C/c = p_{12}(T'_C - X'_C/c) = p_{13}(T''_C)$$

Hence, since $X'_C = c\phi_{23}(T'_C)$,

$$T'_C + \phi_{23}(T'_C) = p_{12}p_{31}(T''_C), \tag{56}$$
$$T'_C - \phi_{23}(T'_C) = p_{21}p_{13}(T''_C) \tag{57}$$

These two equations, by elimination of T''_C, determine the function ϕ_{23}, and they then fix T''_C as a function of T'_C, and so determine the function f_{23} Eliminating T''_C we have

$$T'_C + \phi_{23}(T'_C) = p_{12}p_{31}p_{31}p_{12}\{(T'_C - \phi_{23}(T'_C)\}, \tag{58}$$

and this must be an identity true for all T'_C

Consider next events at B. For these, $X'_B = 0$. Hence we have

$$T_B+X_B/c = p_{21}(T'_B) = p_{31}(T''_B+X''_B/c),$$
$$T_B-X_B/c = p_{12}(T'_B) = p_{13}(T''_B-X''_B/c).$$

Hence, since $X''_B = c\phi_{32}(T''_B)$,

$$T''_B+\phi_{32}(T''_B) = p_{13}p_{21}(T'_B), \tag{59}$$
$$T''_B-\phi_{32}(T''_B) = p_{31}p_{12}(T'_B). \tag{60}$$

These two equations, by elimination of T'_B, determine the function ϕ_{32}, and they then fix T'_B as a function of T''_B, and so determine the function f_{32}. Eliminating T'_B we have

$$T''_B+\phi_{32}(T''_B) = p_{13}p_{21}p_{21}p_{13}\{T''_B-\phi_{32}(T''_B)\}, \tag{61}$$

and this must be an identity true for all T''_B.

Now choose $T''_B = T'_C = \zeta$ in these identities, (58) and (61). They may then be written (reversing the second one)

$$\zeta+\phi_{23}(\zeta) = p_{12}p_{31}p_{31}p_{12}\{\zeta-\phi_{23}(\zeta)\},$$
$$\zeta-\phi_{32}(\zeta) = p_{31}p_{12}p_{12}p_{31}\{\zeta+\phi_{32}(\zeta)\}.$$

Since $p_{12}p_{31} = p_{31}p_{12}$, these imply

$$\phi_{23} \equiv -\phi_{32}, \tag{62}$$

which is one of the desired results.

Further, elimination of ϕ_{23} between (56) and (57) gives

$$T'_C = \tfrac{1}{2}[p_{12}p_{31}(T''_C)+p_{21}p_{13}(T''_C)], \tag{63}$$

and elimination of ϕ_{32} between (59) and (60) gives

$$T''_B = \tfrac{1}{2}[p_{13}p_{21}(T'_B)+p_{31}p_{12}(T'_B)].$$

Since $p_{13}p_{21} = p_{21}p_{13}$ and $p_{12}p_{31} = p_{31}p_{12}$, the last equation may be written

$$T''_B = \tfrac{1}{2}[p_{12}p_{31}(T'_B)+p_{21}p_{13}(T'_B)]. \tag{63'}$$

Hence T'_C is the same function of T''_C as T''_B is of T'_B. Hence

$$f_{23} \equiv f_{32}. \tag{64}$$

This is the desired result. Hence $B \equiv C$.

43. We have now proved that the necessary and sufficient conditions, given $A \equiv B$, $A \equiv C$, for the proposition $B \equiv C$ to be true, are

$$p_{12}p_{31} = p_{31}p_{12},$$

which implies

$$p_{13}p_{21} = p_{21}p_{13},$$

and that then the common value of the first is p_{32}, of the second p_{23}.

44 Suppose now that we have any number of observers equivalent in pairs Then for any suffixes m, n, r the following conditions must hold

$$p_{mn} = p_{mr}p_{rn} = p_{rn}p_{mr}, \tag{65}$$

$$p_{nm} = p_{nr}p_{rm} = p_{rm}p_{nr} \tag{65'}$$

It now follows that any two p's commute, whatever their indices For

$$\begin{aligned} p_{mn}p_{rs} &= p_{m\mu}p_{\mu n}p_{r\mu}p_{\mu s} \\ &= p_{\mu n}p_{m\mu}p_{\mu s}p_{r\mu} \\ &= p_{\mu n}p_{ms}p_{r\mu} \\ &= p_{\mu\rho}p_{\rho n}p_{m\rho}p_{\rho s}p_{r\rho}p_{\rho\mu} \\ &= p_{\mu\rho}p_{mn}p_{rs}p_{\rho\mu}. \end{aligned} \tag{66}$$

Now take $\mu = n$, $\rho = m$ The last equation becomes, since $p_{nm}p_{mn} = 1$,

$$\begin{aligned} p_{mn}p_{rs} &= p_{nm}p_{mn}p_{rs}p_{mn} \\ &= p_{rs}p_{mn}. \end{aligned} \tag{67}$$

This proves the commutative property †

45 *Example* Suppose $A \equiv B$ and $A \equiv C$, and that A and B are in uniform relative motion with relative velocity V_{12}, and that A and C are in uniform relative motion with uniform velocity V_{13}

Then we have seen that

$$p_{12}(\xi) = \left(\frac{1-V_{12}/c}{1+V_{12}/c}\right)^{\frac{1}{2}}\xi, \qquad p_{13}(\xi) = \left(\frac{1-V_{13}/c}{1+V_{13}/c}\right)^{\frac{1}{2}}\xi$$

Since each operator p_{12} or p_{13} is simply multiplication by a constant factor, p_{12} and p_{31} commute, as do p_{21} and p_{13} Hence $B \equiv C$. Moreover

$$p_{23} = p_{21}p_{13}$$

and so the operation p_{23} is also multiplication by a constant. Hence B and C are in uniform relative motion If V_{23} is the relative velocity of B and C, we have accordingly

$$\left(\frac{1-V_{23}/c}{1+V_{23}/c}\right)^{\frac{1}{2}} = \left(\frac{1+V_{12}/c}{1-V_{12}/c}\right)^{\frac{1}{2}}\left(\frac{1-V_{13}/c}{1+V_{13}/c}\right)^{\frac{1}{2}} \tag{68}$$

Solving for V_{23} we have

$$V_{23} = \frac{V_{13}-V_{12}}{1-V_{12}V_{13}/c^2} \tag{69}$$

This is a particular case of Einstein's additive law for velocities

† It should be noted that in general the p's are not *linear* operators

A one-dimensional universe of discrete particles

46. Let us now attempt to construct a one-dimensional system of discrete particles satisfying Einstein's cosmological principle We require that all equivalent particle-observers of the system shall describe the system in the same way We shall first suppose that every member of the system is equivalent to every other member, so that the particles are equivalent in pairs We then make each particle-observer's description of the *whole* system coincide with any other particle-observer's description

47. Let the observers be $A_0, A_1,$. Let A_0 observe $A_1, A_2, \quad, A_n$, to possess time-distance functions

$$\phi_{01}, \phi_{02}, \quad, \phi_{0n},$$

The distribution is to be such that A_1 observes a similar distribution If this condition is satisfied when we transfer from A_0 to A_1, it will also be satisfied when we transfer from A_1 to A_2, and so on, and thus, by compounding transformations, it will be satisfied when we transfer directly from A_0 to A_n

A_0 describes the motion of A_n as

$$X_n = c\phi_{0,n}(T_n)$$

The motion of A_n will be described by A_1 as

$$X'_{n-1} = c\phi_{1,n-1}(T'_{n-1}),$$

where X'_{n-1}, T'_{n-1} are the transforms of X_n, T_n on going from A_0 to A_1. Hence

$$T'_{n-1}+X'_{n-1}/c = p_{01}(T_n+X_n/c),$$
$$T'_{n-1}-X'_{n-1}/c = p_{10}(T_n-X_n/c),$$

whence

$$T'_{n-1}+\phi_{1,n-1}(T'_{n-1}) = p_{01}\{T_n+\phi_{0,n}(T_n)\},$$
$$T'_{n-1}-\phi_{1,n-1}(T'_{n-1}) = p_{10}\{T_n-\phi_{0,n}(T_n)\}.$$

But

$$\phi_{1,n-1} \equiv \phi_{0,n-1},$$

since A_1 describes the $(n-1)$th particle from himself in the same terms as A_0 describes the $(n-1)$th particle from himself Thus

$$T'_{n-1}+\phi_{0,n-1}(T'_n) = p_{01}\{T_n+\phi_{0,n}(T_n)\},$$
$$T'_{n-1}-\phi_{0,n-1}(T'_n) = p_{10}\{T_n-\phi_{0,n}(T_n)\}$$

Now for any suffix r and any argument ξ,

$$\xi+\phi_{0,r}(\xi) \equiv p_{r,0}f_{0,r}(\xi),$$
$$\xi-\phi_{0,r}(\xi) \equiv p_{0,r}f_{0,r}(\xi)$$

Hence
$$p_{n-1,0}f_{0,n-1}(T'_{n-1}) = p_{01}p_{n0}f_{0n}(T_n),$$
$$p_{0,n-1}f_{0,n-1}(T'_{n-1}) = p_{10}p_{0n}f_{0n}(T_n)$$
Eliminating T'_{n-1} we have the identity
$$p_{0,n-1}p_{01}p_{n0} = p_{n-1,0}p_{10}p_{0n},$$
or
$$p_{0n}p_{0n} = p_{01}p_{0,n-1}p_{0,n-1}p_{01}$$
A solution of this equation, since the p's commute, is
$$p_{0n} = p_{01}p_{0,n-1},$$
which gives by recurrence
$$p_{0n} = (p_{01})^n.$$
Then
$$p_{mn} = p_{m0}p_{0n} = [(p_{01})^m]^{-1}[p_{01}]^n$$
$$= (p_{01})^{n-m}. \tag{70}$$

This is the complete solution to the problem of finding the p-functions in terms of the p-function connecting A_0 and A_1 The one-dimensional system is now fully determined, given only the relative motion of A_0 and A_1 ($A_1 \equiv A_0$)

48 *Example* In the case of *uniform* relative motion in one dimension, the one-dimensional system satisfying Einstein's cosmological principle is given, by the above formula, by
$$\frac{1-V_{0n}/c}{1+V_{0n}/c} = \left(\frac{1+V_{01}/c}{1-V_{01}/c}\right)^n \tag{71}$$
or
$$V_{0n} = c\frac{(1+V_{01}/c)^n-(1-V_{01}/c)^n}{(1+V_{01}/c)^n+(1-V_{01}/c)^n}, \tag{71'}$$
where V_{0n} is the observed velocity of the nth observer, as seen by A_0

49 The latter result can be derived without using operational methods, as follows We have by Einstein's addition formula (69) (which is readily proved directly),
$$V_{1,n+1} = \frac{V_{0,n+1}-V_{01}}{1-V_{01}V_{0,n+1}/c^2}$$
But
$$V_{1,n+1} = V_{0,n}$$
Hence
$$V_{0,n} = \frac{V_{0,n+1}-V_{01}}{1-V_{01}V_{0,n+1}/c^2},$$
or putting $V_{0,n+1}/c = \alpha_{n+1}$, and solving,
$$\alpha_{n+1} = \frac{\alpha_n+\alpha_1}{1+\alpha_1\alpha_n}$$

Hence
$$1+\alpha_{n+1} = \frac{(1+\alpha_1)(1+\alpha_n)}{1+\alpha_1\,\alpha_n},$$
$$1-\alpha_{n+1} = \frac{(1-\alpha_1)(1-\alpha_n)}{1+\alpha_1\,\alpha_n},$$
whence
$$\frac{1+\alpha_{n+1}}{1-\alpha_{n+1}} = \frac{1+\alpha_1}{1-\alpha_1}\,\frac{1+\alpha_n}{1-\alpha_n}$$
This recurrence relation gives
$$\frac{1+\alpha_n}{1-\alpha_n} = \left(\frac{1+\alpha_1}{1-\alpha_1}\right)^n, \tag{72}$$
which is precisely the relation (71) already found

50. The sequence of velocities $\ldots, V_{0n}, \ldots$ has for its limit as $n \to \infty$ the value c Thus the one-dimensional system of discrete particles described in the same way from every particle includes an infinity of particles, whose velocities tend on each side to the velocity of light for the very distant particles. For n large, the differences in velocity become small and the number of particles δn with velocities u within a given range u, $u+\delta u$ is given by
$$\delta n = \delta u\left[\frac{\partial}{\partial u}\log\frac{1-u/c}{1+u/c}\right]\Big/\log\frac{1-V_{01}/c}{1+V_{01}/c}$$
$$= \frac{\delta u}{c(1-u^2/c^2)}\,\frac{2}{\log(1+V_{01}/c)-\log(1-V_{01}/c)}$$
It appears that the discrete distribution of velocities approximates to a continuous one for n large, for any value of V_{01}, or for all n, for V_{01} sufficiently small

Summary

51. In this chapter the relationships of 'equivalent' particle-observers have been investigated Two particle-observers A and B are equivalent when A can describe the totality of his observations on B in the same way as B can describe the totality of his observations on A When A and B are provided with clocks only as measuring apparatus, the observations they can make are the timing of the sending and reception of light-signals and the reading of one another's clocks, the latter being physically equivalent to Doppler effect measures The observers can ascertain whether they are equivalent by plotting their observations and communicating them to one another, and comparing graphs

It appears that the special theory of relativity, as embodied in the Lorentz formulae connecting A's and B's observations of an event E, is an immediate consequence of the equivalence of two observers A and B who find themselves in 'uniform relative motion', where the latter phrase is carefully defined in terms of clock observations and constructed coordinates The special theory of relativity is an embodiment in symbols of the notion of equivalence For the case of one dimension, the Lorentz formulae have been generalized to equivalent observers in non-uniform relative motion

It is shown that each of two equivalent observers can assign epochs and distances to distant events, by means of clock-measures only, which coincide with the time-assignments and rigid metre-scale measures commonly used in physics The 'constancy of the velocity of light' is a convention, a consequence of the identity of the conventions used by the two observers for constructing coordinates out of clock-observations, assuming as an observational basis that it is possible to realize in nature 'equivalent' observers The totality of events to which an observer assigns the same time-coordinate, as conventionally defined, form a 'world-wide instant' for this observer, and such events are simultaneous for this observer, in the sense used in physics and commonly used in ordinary life The symbols t, t' occurring in the descriptions of an event E by A and B may be regarded as denoting the Newtonian times of the event in the experiences of the two observers concerned 'Simultaneity' is a concept applicable to a pair of events and an observer The interval of time between two events in different places, in the experience of a given observer, is the difference between their time-coordinates as measured by observations with the clock of the observer concerned.

SUPPLEMENTARY NOTES

P 48 It is to be particularly noted that the phrase 'the length of a rigid rod' has no meaning until we have defined simultaneity, and accordingly has different meanings depending on the observer in whose experience the simultaneity is reckoned

P 49 It should be noted that we appear here to be using 'mixed' measures to ascertain x and y, since A in equations (46) is making use of certain observations by B, but it is easily verified that A will obtain the same values for x and y if he measures the range R of P by means of t_1 and t_4, measures the apparent angular position of P, and then uses the cosine and sine of the angle BAP to give him x and y

III

THE COSMOLOGICAL PRINCIPLE

52 WE have already defined a system of particles satisfying Einstein's cosmological principle as such that if A and B are two equivalent particles of the system, then the description of the system by A (in terms of his clock-measures or associated coordinates) coincides with the description of the system by B in terms of B's clock-measures or associated coordinates In this chapter we consider briefly why such systems are proposed for comparison with the universe disclosed in astronomy

53 We shall throughout use the cosmological principle solely as a principle of selection determining what systems are to be the objects of our study Any system of gravitating particles may be considered as forming a 'universe' For example, we might consider a single 'massive' particle, and examine the behaviour of small test-particles allowed to move freely in its presence We can explore the 'position' occupied by such a free particle at various 'epochs' reckoned by an observer and his clock situated on the massive particle, and its inferred velocities and accelerations The gravitational field of the 'massive' particle is simply a description of the aggregate of accelerations undergone by a test-particle at all points of the space constructed by the observer on the 'massive' particle By calling the given particle 'massive', we mean here simply that the universe so constructed has a centre, namely the massive particle itself, and that an objective, observational meaning attaches to the distance of any point of the space from the massive particle. By a 'small' test-particle we mean a free particle whose presence has no 'influence' on the motion of any other free particle. Such a system of one massive particle and one free particle is usually called a Keplerian system

Or we may consider a system consisting of two massive particles, and analyse their relative motion When the two particles are 'equivalent' in the sense here used, we should describe the particles as of equal mass, when they are not equivalent, of unequal mass Such a system may be called the two-body or double-star problem We may consider similarly three-body problems, and problems of any higher degree of complexity.

54. We may consider also continuous distributions of matter For example we may consider a continuous distribution of matter in rotation about an axis with constant angular velocity This, a famous classical object of study, may be called the rotational problem We may consider further a continuous distribution of matter in the presence of a single external massive particle This is the tidal problem We may combine the rotational, tidal, and double-star problems, and consider two rotating continuous distributions in one another's presence Lastly, we may consider any number of rotating continuous distributions of matter in one another's presence

55. All such systems may be taken as constituting universes. In general, every system of this kind possesses preferential points or axes of reference The Keplerian problem contains a preferential origin The double-star problem contains two preferential origins The tidal problem may also contain two preferential origins. The rotational problem contains a preferential origin and a preferential axis. Systems could be considered containing preferential planes, or preferential directions of density-gradient In all such systems the preferential geometrical element could be recognized and located by means of actual observations conducted on the system

56 In addition to systems containing preferential particles or preferential geometrical elements, it might be possible to construct systems containing no preferential particles or preferential geometrical elements These, if they exist, will form a distinct branch of gravitational studies, worthy of examination on their own merits and on account of their own intrinsic interest

The question arises, if we wish to consider the gravitational field of the whole universe, what system are we to propose for examination? All that the mathematician can do is to ascertain the properties of proposed systems, describe the manner in which they would be presented to observation, and leave them for comparison with the results of astronomical investigation. But the mathematician requires some guide as to which systems he should first consider.

57. The spiral or extra-galactic nebulae, which appear to be the bricks out of which the universe is constructed, may be represented to a first approximation by particles And a volume containing a not

too small number of such particles may be represented as filled with a quasi-continuous distribution of such particles This procedure is fully comparable with the representation of a solid or liquid, composed of molecules and atoms, by a continuous density-distribution It is called a macroscopic procedure In ordinary physics, macroscopic procedure differs from microscopic procedure in that the individual molecules and atoms are as it were smoothed out, leaving an average distribution of matter whose average properties will correspond to the grosser observable properties of the solid or liquid or gas concerned In astronomical physics, microscopic procedure would involve the consideration of the nebulae as units possessing each a structure One possible macroscopic procedure is to smooth out the material forming the nebulae into a quasi-continuous distribution of matter, such that the velocity, density, etc at any point in the smoothed-out distribution represent the average velocity, density, etc in a typical volume element, centred at that point, containing a not too small number of nebulae Such a distribution may be called a 'hydrodynamical' representation of the universe—hydrodynamical because, in classical, non-turbulent hydrodynamics, there is a definite correlation between position and velocity such that at each point there is a unique velocity Most schemes of representation of the universe hitherto proposed have possessed this hydrodynamical character; we shall later consider representations of a much more general kind, which may be called statistical, and which bear the same relation to the hydrodynamic mode of representation as a gas, considered molecularly, bears to a liquid, considered macroscopically For the moment we confine attention to hydrodynamical representations.

A hydrodynamical representation might be chosen of any of the types considered above It might be chosen to possess a single density-maximum, in analogy with the Keplerian problem, and either to have or not to have rotation about an axis, in analogy with the rotation problem It might have two density-maxima in motion, in analogy with the double-star problem, or again many density-maxima, or one density-maximum producing tidal distortions in another region of relatively high density But more fundamental than any of these would be, *prima facie*, a *homogeneous* distribution of matter The behaviour of a homogeneous distribution of matter, possibly in motion, might be regarded as a standard of comparison with observed phenomena

58 It is *a priori* fully possible that the world, smoothed out, would not be homogeneous in any sense which might be attached to the word It is possible, *a priori*, that the world might have a unique density-maximum in it In that case it should be possible to determine the location of this density-maximum by observation, and at every other point there should be a unique direction pointing to the position of the density-maximum, and the motion of the matter near the density-maximum could be regarded as affording a unique standard, which might be taken to define 'rest' The behaviour of matter, laws of nature, and so on in such a system might well depend upon relations involving the distance from this preferential point, the direction towards the preferential point, and the velocity with reference to this preferential point. We should then have no right to believe in the principle of the uniformity of nature as applied to phenomena at different places and involving different directions, nor would the statement of laws of nature be independent of the frames of reference chosen, so that we should have no guarantee or expectation that the principle of relativity would be satisfied

Again, the world might contain two density-maxima, or a finite number of density-maxima In that case we could define a centre of mass or of position in the experience of any one observer, there would be one density-maximum or condensation which has the greatest distance-coordinate from the observer, and this member would be on the outside of the system to this observer If the observer moved towards it, it would either remain on the outside or new condensations would come into view beyond it, and we should have to discuss in the former case an accessible 'edge' or 'rim' to the system, in the latter case whether the 'new' condensations were fresh creations or old condensations seen in new positions.

I am well aware that some mathematicians believe that such difficulties are at once swept away if we use the concept of 'curved space' I have examined such attempts at explanation with the greatest care, and I have found that in all cases the proposed explanations break down at some point Two-dimensional analogies with hypothetical inhabitants on the surface of a sphere fail as soon as we recall that a survey of the astronomical universe is made by taking a photograph with a telescope and camera, and that, for a telescope of arbitrarily large light-gathering power, either the number of nebulae that can be counted is finite and therefore contains one

faintest and so presumably most distant member, or it is infinite, in which case either the same nebula is photographed as an infinite number of separate spots or the total number of actual nebulae in existence is infinite The latter will be our eventual conclusion Here I am only concerned to argue that the phrase 'curvature of space' used in connexion with astronomical photographs merely involves a mist of mysticism Such photographs can always be interpreted in flat space, and then the assumption of a finite number of density-maxima inevitably leads to some kind of accessible edge of the universe

59 Instead, then, of positing one, two, or any finite number of density-maxima as systems for consideration, the obvious suggestion is to consider a strictly homogeneous density-distribution, if that be possible. Unfortunately the concept of homogeneity breaks down at the outset as soon as we recognize the possibility of the presence of *motion* in the system. Let us examine the concept of homogeneity

Homogeneity

60. Our ordinary definition of homogeneity is as follows Let P, Q be two different particles in a homogeneous distribution Then the density-distribution ρ is homogeneous if, for any pair P, Q, the density at P is equal to the density at Q, $\rho(P) = \rho(Q)$ If the density $\rho(P)$ at P is constant in time in the experience of an observer at P, for all P, this definition of homogeneity is unambiguous and unassailable A *static* homogeneous distribution is therefore a descriptive possibility, and the investigation, if on a dynamical basis, would turn on the determination of a set of forces which would give a state of equilibrium at each particle

For free particles, the gravitational field would then have to reduce to zero at each particle This is incompatible with the Newtonian mechanics of a system of free particles For if V is the gravitational potential, the gradients $\partial V/\partial x$, $\partial V/\partial y$, $\partial V/\partial z$ would have to be everywhere zero, and accordingly the second derivatives would be zero But the sum of the second derivatives must equal $4\pi\gamma\rho$, which is non-zero save in the case of an 'empty' universe, $\rho = 0$

To avoid this difficulty Einstein in a classical paper introduced a partly empirical correction to the Newtonian law of gravitation, and succeeded in constructing, in a finite curved space, a static model for

the universe This is in fact incompatible with the observed existence of *motion* in the astronomical universe, but it could have been rejected *a priori* as involving an absolute standard of rest in the universe, which would be a fundamental irrationality it would be for ever impossible to explain how the system of particles lighted on this absolute standard of rest, how they 'recognized' it, and indeed the whole state of affairs would be tantamount to the postulate of a stagnant ether Further, the interpretation of the static system in terms of observations that could be actually carried out leads to the most appalling difficulties—it would lead to the conclusion that we should be able to photograph ourselves (our own galaxy) as a distant luminous patch, at rest relative to ourselves and therefore, when plotted in flat space, at a finite accessible point The static universe of Einstein is in fact incompatible with the arbitrariness of an adopted geometry, which is our present standpoint The static universe of Einstein could be discussed in much greater detail, but further discussion would but accentuate its difficulties, quite apart from its non-correspondence with observation The notion of an empirical correction to Newton's law of gravitation, perfectly conceivable when gravitation is treated as an empirical fact of observation, is in any case quite out of place in a scheme of thought which claims to present gravitation as an inherent constituent of reality as disclosed in observation A fundamental discussion of gravitational phenomena in the universe at large, such as we here later attempt, must leave no room for empirical elements Einstein has, in fact, in later writings, rejected utterly any adherence to a belief in the existence of a 'cosmical' constant of the nature of an empirical correction to the law of gravitation, though it must be stated that his law of gravitation in its unmodified form is still far from inevitable

61 Static homogeneous systems being rejected, it is natural to inquire whether progress can be made by considering a non-static homogeneous system The density ρ at P may now vary with the time, but to measure the time we must introduce an observer Consider then an observer O, situated on some particle of the system, who measures the density at P, at an event E at P to which O assigns the epoch t, as $\rho = \rho(P, t)$ Then the system will be homogeneous in the experience of O if the density at P at epoch t is equal to the density at Q at the same epoch t, where P and Q are two arbitrary points

Thus for O, $\rho(P,t) = \rho(Q,t)$ Let now some other observer, O', be situated on some other particle of the system and therefore in general in motion relative to O Then the two events E_P, E_Q at P and Q at time t will be in general assigned different epochs t'_P, t'_Q by O' Hence O' will find the densities at P and Q to be equal at different epochs, and hence in general, since the density is changing at each point, different at P and Q at the *same* epoch in O''s experience Thus in general, if the system is homogeneous in O's experience it will not be homogeneous in O''s experience Thus no unambiguous definition of homogeneity is possible on these lines The point is that as soon as we recognize the impossibility of an objective simultaneity of two events at two different places to two different observers, we are forced also to recognize the impossibility of an objective definition of homogeneity the same for all observers Equality of density is not an attribute of a pair of points or particles alone, but an attribute of a pair of points or particles considered in conjunction with a specified observer, and if it holds for P, Q, and O, it does not hold in general for P, Q, and O'

62. The methods of current relativistic cosmology overcome this difficulty in the following way Let τ be the time measured by the clock carried with the observer at P, and let $\rho = \rho(P,\tau)$ be the density at P. Similarly let $\rho = \rho(Q,\tau)$ be the density at Q, where τ is measured by the clock at Q Then the system is said to be homogeneous if for any two points P, Q, $\rho(P,\tau) = \rho(Q,\tau)$ In this case, if we take two other observers O and O', situated on particles of the system, the events τ will be in general non-simultaneous in the experiences of O and O', and so in their experiences the densities at P and Q will be equal at *different* epochs Hence in general a system homogeneous according to the above definition will not be homogeneous in the experience of O or in the experience of O' In other words the system homogeneous in the above conventional way will not be homogeneous in the sense of ordinary physics, either to O or to O' Ordinary physics, as shown in Chapter II, employs the time t, not τ, to judge simultaneity, it reckons the epochs of events by a single clock carried by a single observer The observers O and O' will not recognize, or analyse, the system open to their observations as homogeneous

This is not a criticism of the definition of homogeneity used in

relativistic cosmology. I merely point out a consequence of it. But it has the disadvantage that the nature of this conventional homogeneity is only definite when the motion of the particles is first prescribed. For only so are the clock-behaviours of the particles at P and Q prescribed. The definition implies that 'identical' clocks have been set up at P and Q, and moreover that they have been synchronized as to their time zeros, and therefore implies either that they were once at relative rest or that they once all coincided in position. It requires, moreover, that P and Q are 'equivalent' particles. Relativistic cosmology begins in fact by first prescribing a set of motions. It does this by selecting a metric, and selecting in this metric a certain family of geodesics to represent the equivalent or fundamental particles. It then ascertains the corresponding density from the 'field equations' of gravitation. There is enormous choice possible in this selection of a metric, and in the selection of geodesics, and there is choice, too, of the field equations, whether with or without the cosmical constant λ. Thus the density implied by a given metric, even after it has been selected, is not completely determinate. The method selects such metrics as will yield a density ρ such that $\rho(P, \tau)$ is independent of P and a function of τ only. This restricts the choice to some extent, but even so leaves an enormous variety of solutions available. Amongst them it cannot distinguish. The metric selected may correspond to space of positive or negative curvature of any magnitude, or of zero curvature. Moreover it cannot directly compare the implied observational phenomena, for the spaces are all different, a separate space is chosen for each proposed system, and troublesome intermediate calculations are required to pass from the conceptual geometrical situation to the implied phenomena. But the fundamental objection to, or disadvantage of, the methods of relativistic cosmology is that they begin by guessing the answer. In the language of general relativity, the problem is to ascertain a metric suitable for representing a system which is homogeneous in the conventional sense defined above. Actually, though the choice of metrics can be shown to be restricted, every investigator must perforce begin by choosing some metric, that is, by assuming the answer. Moreover such a guess is fundamentally of kinematic character, the investigator prescribes a set of geodesics, a set of timed particle-trajectories, and afterwards uses some dynamics, some set of field equations, to yield the corresponding density.

63. Instead of proceeding in this manner, it is surely much more fundamental to start with some general principle The principle must select a class of systems for discussion, and it must replace the attempt to start with homogeneity, which has been shown to break down

64. The attempt to start with homogeneity breaks down simply because of the relativity of time-assignments for different observers and the non-existence of an absolute simultaneity, the same for all observers This is a consequence of the properties of equivalent particles. The natural suggestion is to replace the non-available concept of an objective homogeneity by the concept of the 'equivalence' of all the particles of the system, in the sense defined in Chapter II But this does not suffice to determine the system We have seen from the one-dimensional example of Chapter II that systems of equivalent observers may be constructed in an infinite variety of manners, by the choice of commuting operator-functions p_{mn} Thus the concept of the equivalence of all the particles present in the system does not fully replace homogeneity as giving a determinate system for discussion We gain our object by making each particle-observer not only equivalent separately to each other particle-observer, but by making each particle-observer equivalent to each other in relation to his observations of the *whole* system The notation $A \equiv B$ implies that the totality of A's observations on B are described by A, using his clock, in the same way as the totality of B's observations on A are described by B, using B's clock The notation $A \mathrel{\underline{\equiv}} B$ will be used to imply that the totality of A's observations not only on B but *on the whole system* are described by A in the same way as the totality of B's observations *on the whole system* are described by B, where A, B are any two particle-observers of the system Such a system is a sub-class of the systems satisfying Einstein's cosmological principle The whole class of such systems is defined by the proposition if A and B are two particles of the system such that $A \equiv B$, then also $A \mathrel{\underline{\equiv}} B$ In the former case, *any* two particles are equivalent, so that for all A, B, $A \equiv B$ and $A \mathrel{\underline{\equiv}} B$ In the latter case, $A \mathrel{\underline{\equiv}} B$ is true only for particles satisfying $A \equiv B$ We confine attention for the time being to the former case

Thus Einstein's cosmological principle is to be taken as a definition replacing the unworkable definition of homogeneity, and selecting a class of systems for consideration Whether systems can be con-

structed satisfying Einstein's cosmological principle remains for mathematical investigation Whether, when they have been constructed mathematically, they agree in their observable properties with the universe of nature remains for astronomical investigation The principle appears, however, to be the simplest general principle which has not the obvious objection of beginning by inserting preferential view-points, frames of reference, a centre or centres of symmetry, axes of symmetry, planes of symmetry, all of which would render doubtful the application of the principle of the uniformity of nature and the principle of relativity The principle prescribes a norm of behaviour, against which actual astronomical observations may be projected

65. It must particularly be noted that the principle does not mean that if A and B observe a third particle P, then A describes P in the same way as B describes P It means that if A's experiences contain a P, then B's experiences contain a Q described by B in the same way as A describes P A's and B's pictures are superposable, but the superposition is not one of identical particles but of corresponding particles This will become clearer in later examples, an instance of it has already occurred in the one-dimensional system of discrete particles considered at the end of Chapter II

66. Whether the universe may be *expected* to be representable by a system satisfying the cosmological principle is a metaphysical question To discuss it in detail would be outside the plan of this book We have already suggested that the system of the smoothed-out universe, if it violated the cosmological principle, would create grave difficulties, of the type of the existence of an absolute standard of rest, for example My own private opinion is that the universe must satisfy the cosmological principle, because it would be impossible for an act of creation to be possible which would result in anything else To create, for example, a universe consisting of only two particles would involve a fixation of the 'position' and 'velocity' of their centre of mass, in the desert of otherwise-emptiness which space would be The featurelessness of space, the non-existence of any frame of reference with regard to which the 'position' and 'velocity' of the centre of mass could be specified, would preclude the description of the system which had been created With God all things are not possible The individual mathematician can discuss a universe

of two particles because he supplies in thought a frame of reference not available in actual nature. Even in the act of creation, God requires, for certain bricks, straw. For similar reasons I do not believe that the universe could consist of any finite number of particles, similar difficulties would arise. The systems which the cosmological principle selects for our consideration will always be found to contain an infinity of particles, accessible to observation. The systems which relativistic cosmology constructs always contain, at any given epoch of observation, a finite number of particles accessible to observation, but create further particles as the epoch of observation advances. (This will be established later.) This to my mind involves the rejection of the schemes of relativistic cosmology, both on account of their finite population of particles at any one epoch of observation, and their requirement of progressive creation.

67. Such metaphysical questions will not be pursued for the present. We shall be content to construct systems satisfying the cosmological principle, explore their properties, and compare them both with observation and with the schemes proposed in current relativistic cosmology, always regarding them as the simplest world-schemes that could be suggested. They are as simple in relation to the study of the cosmological problem—the problem of the distribution of matter and motion in the universe—as, for example, the study of a 'homogeneous' specimen of *silver* is in relation to silver in general. Any given specimen of silver may not be homogeneous—in the sense of course of ordinary physics—but we want to tabulate the properties of homogeneous silver as a standard of comparison. We have only one universe to observe, but we still want the equivalent of a homogeneous universe as a standard of comparison, and such standards are provided by systems satisfying the cosmological principle.

68. Our analysis of the notion of homogeneity broke down owing to the impossibility of defining an objective simultaneity. We saw that one way of reintroducing simultaneity, the way adopted in relativistic cosmology, was to use the 'cosmic' time τ, the time kept by the clock moving with the particle near which the density was to be evaluated. Another way is to *ignore* the failure of objective simultaneity, and to use a Newtonian time t supposed to be the same for all observers. Such a procedure is admittedly not logical or self-consistent. Nevertheless it suggests the problem of attempting to construct a system

satisfying the cosmological principle (now suitably modified) within the domain of Newtonian mechanics and gravitation, using Newtonian relativity, i e the Newtonian rules of transformation from one observer's Newtonian space and time to another observer's space and time This problem has recently been solved,† and the solution is given in Chapter XVI It proves to be of no mere academic interest, for it leads to a scheme of conceptual equations identical with the conceptual equations of relativistic cosmology The symbols occurring in the equations are related in identical ways It is only the code of interpretation of these symbols in terms of observations—the code for passing from symbols to sense-data—which differs This is remarkable, as it is often considered that all we know about the universe is contained in differential equations of certain forms Actually it is the *interpretation* of the symbols which is the significant thing The identity of the equations derived in Newtonian fashion with those derived in relativistic fashion is the more remarkable in that it discloses a deeper parallelism between the two schemes of thought, in this the most fundamental of all gravitational problems, than has hitherto been realized No sense can be attached to saying that the one is more accurate than the other Each is perfect and exact in its own domain It is impossible to tell, by looking at the fundamental equations, whether they have been derived by the rules of the Newtonian scheme or by the rules of the relativistic scheme The sole question is the deciding as to what set of rules shall be used for predicting observations from the equations I believe that this has a deeper significance than has yet been explored.

69. In constructing systems satisfying the cosmological principle, I shall in the first place confine attention to the case in which A and B (any two particles) are not only equivalent but are in *uniform* relative motion I shall show in due course that the circumstances are such that they will actually move freely with uniform relative velocity, so that the motion is a natural or gravitational motion Extensions to relatively accelerated equivalent observers could be undertaken as soon as the formal problem of obtaining generalized Lorentz formulae in three dimensions has been solved. It is by no means certain that such extensions are possible, or that if

† Partly by the writer alone, partly by the writer and Dr W H McCrea in collaboration (*Quart J of Math* (*Oxford*), **5**, 64 and 73, March 1934)

possible they will yield anything fundamentally new Actually the results obtained with restriction to the case of uniform relative motion have exactly the same degree of generality (same multiplicity of arbitrary constants) as the complete set of results given by relativistic cosmology They therefore cover a corresponding ground, in simpler fashion The results are in any case so rich, the methods so novel, and the conclusions so extraordinary that the special case of uniform relative motion of equivalent observers, if special case it be, fully justifies detailed consideration, and requires to be discussed first before any progess would be likely to be possible with a calculus of greater generality

In Part III we construct systems satisfying the cosmological principle in the wider sense, i e systems containing other particles than the equivalent particles, but still satisfying the condition that for all pairs A, B for which $A \equiv B$, then $A \equiv B$

PART II

KINEMATIC WORLD-MODELS

IV

ELEMENTARY KINEMATIC CONSIDERATIONS

70. To gain insight into the properties of systems satisfying the cosmological principle, let us first ignore the relativity of time-determinations and suppose that it is legitimate to adopt Newtonian time the same for all observers Since such time-determinations coincide with actual time-determinations by equivalent observers in the immediate neighbourhood of the observers concerned, the results we obtain will in any case hold good locally

71. Consider then a system of equivalent particles P, in motion in some manner, and satisfying the cosmological principle Each particle-observer is to adopt Newtonian time and Newtonian space Fix attention on any one particle-observer O of the system and let him observe the velocities relative to himself of all other particles P Each particle, at any arbitrary time t, will possess a position-vector $\mathbf{r}$ with respect to O, and a velocity-vector $\mathbf{v}$ Observer O can make a complete enumeration of the velocities $\mathbf{v}$ of particles P, cataloguing them in terms of their position-vectors $\mathbf{r}$ at all times t He can summarize his catalogue in the form of a formula

$$\mathbf{v} = f(\mathbf{r}, t), \tag{1}$$

where the vector-function f is simply a function describing all the velocities concerned

72 Now take a second observer O' situated on some other one of the particles P Since the system is given to satisfy the cosmological principle, the description of the velocity-behaviour of the system by O coincides with the description by O'. Hence the velocity $\mathbf{v}'$ relative to O' of any particle whose position-vector with respect to O' is $\mathbf{r}'$ is given by

$$\mathbf{v}' = f(\mathbf{r}', t), \tag{1'}$$

where f is the same vector-function as before Now let $\mathbf{r}_0$ be the position-vector of O' with respect to O. Since O' is situated on a particle of the system, the velocity $\mathbf{v}_0$ of O' relative to O is given by

$$\mathbf{v}_0 = f(\mathbf{r}_0, t) \tag{2}$$

But by Newtonian relativity

$$\mathbf{v}' = \mathbf{v} - \mathbf{v}_0, \tag{3}$$

$$\mathbf{r}' = \mathbf{r} - \mathbf{r}_0 \tag{3'}$$

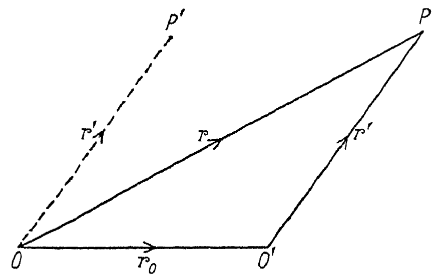


FIG 7 Enumeration of velocities by two observers O, O' such that $O \equiv O'$

Hence $$f(\mathbf{r}-\mathbf{r}_0, t) = f(\mathbf{r}, t) - f(\mathbf{r}_0, t) \tag{4}$$

Hence f is a linear vector-function of the vector $\mathbf{r}$, and its most general expression is of the form of the inner product

$$f = T(t)\ \mathbf{r}, \tag{5}$$

where $T(t)$ is a Cartesian tensor (of the second rank) The components of the tensor $T(t)$ in any frame of reference associated with O must be independent of the frame of reference chosen, and so the same in all frames of reference It is therefore an isotropic tensor† of rank 2, and so proportional to the idem-tensor U, of the form $\mathbf{ii}+\mathbf{jj}+\mathbf{kk}$ in any triad of reference $(\mathbf{i}, \mathbf{j}, \mathbf{k})$, where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are mutually perpendicular unit vectors Here $T(t) = F(t)U$ and $T(t)\ \mathbf{r} = F(t)U\ \mathbf{r} = F(t)\mathbf{r}$ Thus

$$\mathbf{v} = \mathbf{r}F(t) \tag{6}$$

In words, the velocity-vector of P is along and proportional to the position-vector OP, for all origins of reference O

Thus in a system satisfying Einstein's cosmological principle, from kinematic considerations alone, without any appeal to dynamics, we establish the radial character of the motion, and the proportionality of velocity to radial distance So far our result holds only locally

73. This is a well-known characteristic of the motions of the extra-galactic nebulae as judged by the Doppler effects in their spectra It is known as Hubble's law, since it was discovered empirically from observational data by Dr Edwin Hubble, of the Mount Wilson

† Cf H Jeffreys, *Cartesian Tensors*

Observatory Actually the Doppler shifts in the nebular spectra are directed towards the *red*, so that the inferred velocities are *outward* velocities The above argument does not show why the velocities are all outward, as the function $F(t)$ may be negative as far as the foregoing deduction goes The fact that Hubble's law is obeyed by the extra-galactic nebulae is a strong piece of evidence in favour of the hypothesis that the smoothed-out universe, in which the nebulae are represented as particles, is a system satisfying the cosmological principle, and is a sufficient justification for the further study of such systems It will be noticed that our derivation of $\mathbf{v} = \mathbf{r}F(t)$ is quite independent of gravitational considerations, and holds good locally on any theory or law of gravitation whatever It is the most primitive property of a system satisfying the cosmological principle

74. Comparison of the observed velocities of the extra-galactic nebulae with their observed distances gives at once the present numerical value of the function $F(t)$ If observations made at a succession of times t showed an appreciable alteration in the values of $F(t)$, we should be enabled at once to evaluate the accelerations of the nebulae For, differentiating 'following the motion' of any particle, we have

$$\frac{d\mathbf{v}}{dt} = \mathbf{v}F(t)+\mathbf{r}F'(t)$$
$$= \mathbf{r}[\{F(t)\}^2+F'(t)]$$

Hence the nebulae have positive, zero, or negative accelerations according as

$$F'(t)+\{F(t)\}^2 \gtreqless 0$$

In any range of t in which this expression has a constant sign, this integrates in the form

$$t-t_0 \gtreqless \frac{1}{F(t)},$$

where t_0 is some constant, i e

$$F(t) \gtreqless \frac{1}{t-t_0}$$

Actually observation is not yet sufficiently precise to disclose any day-to-day or year-to-year variation in the observed Doppler shifts Within our present accessible range of time-differences, then, the motions may be considered uniform To this order of accuracy, therefore, $F^2+F' = 0$, and so

$$F(t) = \frac{1}{t-t_0},$$

whence
$$\mathbf{v} = \frac{\mathbf{r}}{t-t_0},$$
and is constant for any one particle This may be written in scalars
$$\frac{1}{r}\frac{dr}{dt} = \frac{1}{t-t_0}$$
or
$$r = v(t-t_0), \tag{7}$$
which is the distance-time relation for any one nebula, v being its constant velocity According to this, at the interval $t-t_0$ ago, r was zero, and the nebulae were all close together—in this idealization coincident The value $t-t_0$ may be evaluated from the observed values of r and v for the nebulae Adopting a linear rate of increase of velocity with distance of 500 km. sec $^{-1}$ per 10^6 parsecs (1 parsec $= 3\ 08\times 10^{13}$ km), the present value of $t-t_0$ is
$$\frac{3\ 08\times 10^{19}}{500} = 0\cdot 6\times 10^{17} \text{ seconds,}$$
or since 1 year $= 3\ 16\times 10^7$ seconds, approximately,
$$t-t_0 = 2\times 10^9 \text{ years}$$
If we reckon t, our present epoch, from the epoch at which according to this evaluation the nebulae were close together, we may take $t_0 = 0$, and then
$$v = \frac{r}{t}, \tag{8}$$
where the 'present' value of t is 2×10^9 years

75. We must later examine under what circumstances the nebulae could be moving in this way, with constant relative velocities As it stands the estimation is a gross extrapolation In the meantime we proceed to a further property of the system of particles satisfying Einstein's cosmological principle

76. We have been adopting Newtonian time for all observers—a sufficiently accurate assumption for observers not too far apart For such observers, as we saw, 'homogeneity' has a meaning, and the system satisfying the cosmological principle will be homogeneous, that is, locally homogeneous Let $n(t)$ be the number of particles present per unit volume, $n(t)$ is independent of position Now since v is fully correlated with r, the motion is of hydrodynamical character Then since the number of particles is conserved, the hydrodynamical

equation of continuity must be satisfied This is, in polar coordinates, since $v = |\mathbf{v}|$ depends for given t only on $r = |\mathbf{r}|$,

$$\frac{\partial n}{\partial t}+\frac{1}{r^2}\frac{\partial}{\partial r}(r^2nv) = 0 \tag{9}$$

Since n is a function of t only, we may write this equation as

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2v) = -\frac{1}{n}\frac{dn}{dt} \tag{10}$$

Here the right-hand side is a function of t only We may write accordingly

$$-\frac{1}{n}\frac{dn}{dt} = 3\phi(t), \tag{11}$$

whence

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2v) = 3\phi(t)$$

Integrating partially with respect to r we have

$$r^2v = r^3\phi(t)+\psi(t)$$

or

$$v = r\phi(t)+\frac{1}{r^2}\psi(t)$$

For this to be consistent with (6)† we must have

$$\psi(t) \equiv 0, \qquad \phi(t) \equiv F(t) \tag{12}$$

Hence the particle-density $n(t)$ varies according to the law

$$\frac{1}{n}\frac{dn}{dt} = -3F(t)$$

or

$$n(t) = Ne^{-3\int_T^t F(t)\,dt}, \tag{13}$$

where N is the particle-density at $t = T$, and is arbitrary This

† Alternatively, $\psi(t) \equiv 0$ follows from the consideration that there can be no singularity at the origin $r = 0$ Hubble's law $v = r\phi(t)$ then follows without recourse to the cosmological principle Thus Hubble's velocity distance proportionality may be considered either as the most primitive property of a system whose velocity distribution, as recorded by different observers not too far apart, satisfies the cosmological principle considered as a definition, or it may be considered as the most primitive property, namely an immediate consequence of the condition of hydrodynamical continuity, of a system of matter in motion which is locally homogeneous This shows how inevitable is the distance velocity proportionality It is remarkable that neither of these simple derivations has yet found a place in the standard treatments of the subject based on relativistic cosmology Instead, in those treatments, the proportionality emerges as a mere consequence of mathematical juggling with symbols appertaining to curved spaces It is in fact almost always true that the complicated mathematical machinery of general relativity obscures by its very power the inner nature of the phenomena it attempts to explore. In this as in other examples, far more insight is gained by elementary considerations

determines the density-behaviour in terms of the function $F(t)$, itself given from the velocity-distance proportionality If $F(t) > 0$, as is implied by the observed recession, $n(t)$ must be decreasing as t increases Thus the system is in a state of continuous *dilution* This is a kinematic consequence of the expansion For $F(t) = 1/t$, $n(t) = NT^3/t^3$

The world-picture

77. We pause here to make an important inference from these kinematic relations, which are necessarily true locally Suppose that a photograph is taken, at time t, of a region of the heavens Let r_1 be the apparent distance of any one nebula in this photograph Then the light recorded in the photograph left the nebula at epoch $t_1 = t - r_1/c$, and since t is the same for all the nebulae photographed, t_1 is different for nebulae with different values of r_1 The observed distances are less than the *present* distances, owing to the recession Hence if the distribution is actually locally homogeneous (at time t), the distribution in the photograph will be compressed radially, the compression increasing with increasing distance This means that in the photograph, the density must increase outwards, and the distribution as photographed will not appear homogeneous This will be found to be a very large effect

Suppose the nebulae inside a given solid angle $d\omega$ and within a given range (r_1, r_1+dr_1) are counted on a photograph taken at time t Let the number be $\nu r_1^2\, dr_1\, d\omega$ If the distribution as photographed appeared homogeneous, ν would be independent of r_1 Actually it can be shown,† as a consequence merely of the kinematic formulae obtained above, that the number ν, which of course must also depend on t, depends on r_1 according to the first-order formula

$$\nu = \nu(r_1, t) = n(t)\left(1+4\frac{v_1}{c}\right), \tag{14}$$

where v_1 is the observed velocity af the nebula at distance r_1, and $n(t)$ is the particle-density at time t close to the observer, i e. the value of ν at time t for $r_1 = 0$

This is a very important formula Owing to the large factor $1+4\frac{v_1}{c}$, the effect is large even for small values of v_1/c, and it should be within the present range of observation Actually Hubble has found

† See Appendix, Note 1.

evidence of an outward increase of density from the photographic counts, when allowances of various kinds are made for the diminution of nebular brightness, as observed, due to recession of the source, and consequent displacement of the whole continuous spectrum to the red Formula (14) above is an immediate consequence of the velocity-distance proportionality if red-shifts are interpreted as actual motions of recession, and its verification would constitute first-hand evidence that the nebulae are in actual motion outwards We have already seen that if the nebulae can be represented as particles 'equivalent' to one another in the technical sense, then Doppler shifts must necessarily be interpreted as actual motions in the ordinary sense used in physics The verification of the above formula would then be a piece of evidence in favour of the 'equivalence' of the nebulae, and so of the obedience of the system of the nebulae to the cosmological principle

The kinematics of a cluster of particles in uniform relative motion

78. As far as the preceding considerations go, the motion of a system obeying locally the cosmological principle might be either one of approach or one of recession Let us see whether there are any considerations distinguishing between the two

We have seen that present observations cannot detect the accelerations, if any, of the nebulae Let us then examine the particular case of a system of particles *constrained* to move with uniform relative velocities, without imposing the restriction that it is to satisfy the cosmological principle

79 Consider a cluster of such particles, of any space- and velocity-distributions, occupying at a given moment the interior of a domain S and surrounded by 'empty space' For simplicity we shall take S to be a sphere, but the arguments are perfectly general Let O be a particle of the system near the centre of the sphere Consider the particles in the neighbourhood of some point P of the cluster The outward-moving particles will soon penetrate into the 'empty space' near P, the faster ones getting ahead and forming an expanding frontier The inward-moving particles will travel along chords of the sphere, approach to a minimum distance from O, and then be reckoned as outward-moving particles, the faster ones again ahead After the lapse of any considerable time, only the very slow-

moving particles will have motions of approach towards O; the remainder will have predominantly motions of recession, approximately radially outwards from O Ultimately recession prevails for all particles The same thing occurs if the system is viewed from any other particle O'

Let $t = t_0$ be the instant when the system is first *given*, in the experience of the observer O. Then if r_0 is the original radius of the

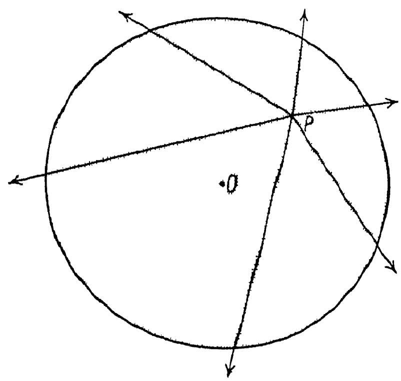


FIG 8 Ultimate recession of all particles forming a cluster when in uniform relative motion.

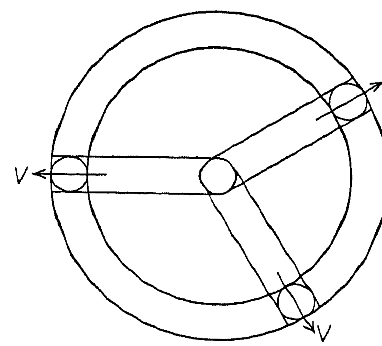


FIG 9 Determination of velocity-distance law

cluster, i e the radius at time t_0, at any later time t the distance r from O of a particle moving with velocity v lies between $v(t-t_0)+r_0$ and $v(t-t_0)-r_0$ Thus

$$v(t-t_0)-r_0 < r < v(t-t_0)+r_0$$

or

$$\frac{r-r_0}{t-t_0} < v < \frac{r+r_0}{t-t_0}. \tag{15}$$

For $t-t_0$ sufficiently large, we have accordingly

$$v \sim \frac{r}{t-t_0} \tag{16}$$

Thus velocity is ultimately correlated with distance The fastest particles will be found in a zone on the outside of the expanding swarm, followed by and partially mixed with the next fastest, and so on However random the original velocity distribution, a sorting-out process occurs, the faster- and slower-moving particles segregating themselves naturally. Moreover, at any given epoch t, the velocities at different distances r tend to become simply proportional to r, according to the above equation

Thus a cluster of particles surrounded by empty space and constrained to move with uniform velocities exhibits an expansion phenomenon, and after a sufficiently long time a predominant recession and a correlation of velocity with position according to a velocity-distance proportionality, reckoned with reference to any assigned particle of the system

80. Such a cluster, containing a finite number of particles and surrounded by empty space, does not satisfy the cosmological principle It is not proposed as a model for the universe But its examination has shown that it exhibits many of the characteristic properties of the system of the nebulae It is easy, in fact, to modify the cluster in a simple way so as to obtain a system of uniformly moving particles satisfying the cosmological principle, as we shall see in the next chapter In the meantime, consideration of the phenomena exhibited by such a cluster affords such insight into the origin of the features exhibited by the nebulae that it is worth while to examine it further

81. First we notice that the system tends to assume the hydrodynamic character. Out of chaos we get a systematic distribution of velocity, inevitably The frontier, or indeed any concentric surface, acts as a Maxwell's sorting demon, letting the swifter particles pass ahead of the slower ones This is a consequence of the original density-inequality, any system containing a density-maximum will tend by diffusion to yield a separation of swift particles from slow ones The systems usually considered in thermodynamics are *enclosed* systems, possessing a boundary Here the system avoids the usual consequences of the second law of thermodynamics partly because it is not walled-in. The system passes naturally, of itself, from a state possessing organization in space-distribution but no organization in velocity-distribution to a state possessing less organization in space-distribution and greater organization in velocity-distribution

82 Secondly, let us consider more closely the 'initial' configuration, the configuration at time $t = t_0$ We have seen that, from this moment on, the system in general expands (Very special conditions would have to be satisfied by the velocity-distribution at $t = t_0$ for the system not to expand.) Now imagine that at $t = t_0$ all the velocities are reversed Then by the same argument as before, unless very special conditions are satisfied, the reversed system will also expand

Next suppose that the original system is allowed to expand some little way, and that then all the velocities are reversed Then the system will be observed to contract to its original ($t = t_0$) configuration and then expand again The configuration $t = t_0$ is thus in general the minimum volume of the configuration in the reversed motion, and so also the minimum volume if the forward motion is extrapolated backwards in time from $t = t_0$ The epoch of minimum volume may be called the natural zero of time for the system, $t = 0$, and thus, unless very special conditions are satisfied at $t = t_0$, the epoch $t = t_0$ at which the system is first given coincides with the natural zero of time for the system Since, if a system be viewed, it is necessarily viewed later than the instant $t = t_0$ at which it was *first given*, it follows that any observation performed on a system of particles in uniform relative motion will in the overwhelming majority of cases reveal the system as an expanding system

It is possible, indeed, that the system at $t = t_0$ may possess such a distribution of velocities that it begins to contract, ultimately, of course, expanding again In that case observation would reveal the system as a contracting system, during an interval following $t = t_0$ Such initial configurations will be highly infrequent compared with the configurations for which expansion occurs from $t = t_0$ on It follows that if a system is started off or *given* at $t = t_0$, it is overwhelmingly probable that it will expand

It is of course easy to construct the special distributions which yield contraction We simply have to take a system at random, let it expand a little, and then reverse its velocities, the result is a contracting system It might be thought in consequence that contracting systems were as frequent as expanding ones This is true Nevertheless observation will almost always disclose expansion, since the contracting phases of a system occur before the natural time-zero, whilst the expanding phases occur after it, and an observation, necessarily made after the system is *first given*, is in the enormous majority of cases later also than the natural time-zero

Whether a system is observed to be expanding or contracting depends on whether the epoch of observation t is such that $t > 0$ or $t < 0$. The information we have about t is that $t > t_0$, where t_0 is the epoch at which the system was first given In general $t_0 = 0$, though exceptional cases may occur for which $t_0 < 0$ or $t_0 > 0$ Only when $t_0 < 0$ is it possible for an observation to disclose contraction

83. To be 'first given' is the same thing as 'to be created' It is always possible to trace backwards the configurations assumed by a system created in time to before the epoch of creation, but propositions about such configurations are necessarily non-verifiable propositions, which accordingly have no significant content † We may modify Wittgenstein's aphorism 'The world is everything that is the case', to 'The world is everything that can be observed to be the case' We see that if a system of the kind considered is created, it is enormously probable that the epoch of creation $t = t_0$ coincides with the natural time-zero, or epoch of minimum volume, $t = 0$ It is much easier to create a system for which $t_0 = 0$ than to create a system for which $t_0 \neq 0$ For, to do this, the creator has merely to create a system at random If a large number of systems are imagined created, then for the enormous majority t_0 will be zero. The chances are that a system created at random will be such that $t_0 = 0$, and hence that any observation made on the system will disclose it as expanding The phenomenon of the expansion can only be understood by taking into account a theory of knowledge Contracting systems are not impossible but are highly improbable

84. The above considerations refer to the times and epochs measured by a single observer How these conclusions are modified in the experiences of other observers will be examined when we construct, in the next chapter, a system of uniformly moving particles, *without* constraints, satisfying the cosmological principle Here we notice that the system here considered possesses a non-zero volume at $t = 0$, and permits only a statistical obedience to the recession law. Slow-moving particles, in the vicinity of the 'centre' of the system, may disobey the recession law This is found to be the case in the system of the nebulae, as observed ‡

85 It is not easy to treat relativistically a system of which the minimum volume is not zero The natural tendency of the mathematical technique employed is to construct an idealization in which the minimum volume is zero If we adopt the view that the world,

† That is, non-verifiable propositions about the world of nature have no significant content In geometry, however, all propositions are unverifiable—they relate simply to the entities defined by the axioms

‡ It is not certain that the Andromeda nebula is receding Its radial velocity is negative, but this may be due to the galactic rotation, which is swinging the earth toward the nebula

in the experience of any one observer, was created in time in the temporal experience of that observer, it is still a quite open question whether it originated as a geometrical particle or point-source, or 'radioactive super-atom' as conjectured by Lemaître, or whether it possessed at the epoch of creation a finite volume. The assumption of a non-zero initial volume involves, however, serious difficulties in a relativistic formulation. For, as we shall see, the relativistic formulation implies that the very distant particles or nebulae possess speeds approaching that of light, which has for a consequence that to observers situated on them the whole system must be contracted to zero volume. The events on them, *now* capable (in principle) of being observed by us, are arbitrarily early in their local time-scales, and correspond as closely as we please to the epoch of creation. 'Creation' or 'earliness in time relative to the natural time-zero' is a *present* characteristic (to us) of the most distant members of the system. This will become clear in the later analysis.

86. In this section we have confined attention to systems of particles in *uniform* relative motion. Current relativistic cosmology contemplates world-models in which relative accelerations are present, but it does so only at the cost of introducing a continuum of 'acts of creation' at finite (non-zero) times in the experience of any given observer. It ensures the satisfaction of the cosmological principle (which includes as one property the centrality of every particle in the field of the remainder) only by creating fresh particles in the vicinity of every already created particle as fast as observations can be made from that particle. It seems convenient to leave the examination of systems of equivalent, relatively accelerated particles, until we give an account of current relativistic cosmology. Here we may be content with having gained insight into the characteristic of *expansion*, as observed to hold good, in contrast to that of contraction, from the consideration of a system of particles in *uniform* relative motion. Expansion, though of the same mathematical nature as contraction, and equally possible as far as mathematical analysis goes, is the overwhelmingly probable phenomenon when regard is paid to observation as our only means of knowledge of the external world.

V

THE SIMPLE KINEMATIC WORLD-SYSTEM AND ITS PROPERTIES

87. We have seen that a local velocity-distance proportionality and a local homogeneity are inevitable characteristics of systems satisfying the cosmological principle Local homogeneity is indeed only to be expected, for we have seen that the cosmological principle is the simplest definition we can adopt which can replace the untenable definition of a world-wide homogeneity We have seen also that a system of particles constrained to move with uniform relative velocities exhibits naturally an expansion, possesses a natural time-origin, and tends to assume a hydrodynamical character with a velocity-distance proportionality of the special type $\mathbf{V} \sim \mathbf{r}/t$, where t is measured from the natural time-zero These were two independent investigations Let us attempt to combine them, and to construct if possible a system of particles in uniform relative motion satisfying the cosmological principle

Preliminary lemma Einstein's velocity-addition formulae

88 Let (x,y,z,t), (x',y',z',t') be the coordinates assigned by two observers A and B in uniform relative motion to any event E on a particle P in motion If the two observers possess a relative velocity U in the direction of the x-axis, then by the Lorentz formulae

$$x' = \frac{x-Ut}{(1-U^2/c^2)^{\frac{1}{2}}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t-Ux/c^2}{(1-U^2/c^2)^{\frac{1}{2}}} \tag{1}$$

Consider a neighbouring event on P, of coordinates $(x+dx, \quad ,t+dt)$ and $(x'+dx', \quad ,t'+dt')$ Then we define as the velocity-components u,v,w of P, as judged by A, the differential coefficients

$$u = \frac{dx}{dt}, \qquad v = \frac{dy}{dt}, \qquad w = \frac{dz}{dt}, \tag{2}$$

and the velocity components u', v', w', as judged by B, the differential coefficients

$$u' = \frac{dx'}{dt'}, \qquad v' = \frac{dy'}{dt'}, \qquad w' = \frac{dz'}{dt'} \tag{3}$$

These are what are meant by the term 'velocity' as customarily used in physics, t and t' being always the epochs assigned by A and B to the events at P as judged by their respective clocks

It follows that

$$u' = \frac{dx - U\,dt}{dt - U\,dx/c^2} = \frac{u-U}{1-uU/c^2},$$

$$v' = \frac{dy(1-U^2/c^2)^{\frac{1}{2}}}{dt - U\,dx/c^2} = \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2},$$

$$w' = \frac{dz(1-U^2/c^2)^{\frac{1}{2}}}{dt - U\,dx/c^2} = \frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2} \tag{4}$$

These formulae are in fact more general than formulae (1), for they are independent of the additions of arbitrary constants to the epochs t and t', and so independent of the synchronization of the zeros of A's and B's clocks They could in fact be demonstrated from consideration of Doppler effects only, converted into velocities, on using the equality of the Doppler effects observed by A and B at one another We have already encountered, in § 45, a particular case of the first of formulae (4)

The invariant velocity-distribution

89. Now consider a system of particles constrained to move with uniform velocities in the experience of an observer situated on some assigned member of the system Then they will have uniform velocities in the experience of any other observer situated on any other member of the system

Let A and B be two such observers, moving with uniform relative velocity $(U, 0, 0)$ Let A assign the velocity $\mathbf{V}$, of components u, v, w, to any particle of the system, and let B assign to the same particle the velocity $\mathbf{V}'$, of components u', v', w'

Let A and B enumerate the velocity-statistics of the system, each in his own experience A counts the number of particles with velocity-components lying between u and $u+du$, v and $v+dv$, w and $w+dw$, he finds that this number is independent of the epoch of observation, and he writes it down as

$$f_A(u, v, w)\,du dv dw,$$

assuming that the particles are sufficiently numerous to be treated statistically Similarly, B enumerates the velocity-statistics in his experience as

$$f_B(u', v', w')\,du' dv' dw'$$

The particles counted by A as inside the velocity domain $du dv dw$ surrounding (u, v, w) will be counted by B as inside the domain $du' dv' dw'$ surrounding (u', v', w') Hence

$$f_A(u, v, w)\,du dv dw = f_B(u', v', w')\,du' dv' dw'$$

But by (4),
$$du'dv'dw' = \frac{\partial(u',v',w')}{\partial(u,v,w)}\,dudvdw$$
$$= \frac{(1-U^2/c^2)^2}{(1-uU/c^2)^4}\,dudvdw$$

Hence
$$f_A(u,v,w) = f_B\left(\frac{u-U}{1-uU/c^2}, \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}\right)\frac{(1-U^2/c^2)^2}{(1-uU/c^2)^4}$$

Now let us attempt to construct a system such that A's description of the velocity-statistics coincides with B's, for any A and B situated on particles of the system Their velocity-pictures are to be identical Then the condition for this is that for any U,
$$f_A \equiv f_B,$$
$$\equiv f,$$
say Then we have the following functional equation for f
$$f(u,v,w) \equiv f\left(\frac{u-U}{1-uU/c^2}, \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}\right)\frac{(1-U^2/c^2)^2}{(1-uU/c^2)^4} \qquad (5)$$

This equation must be satisfied for all values of U occurring in the system Two similar equations must be satisfied, derived by replacing $(U,0,0)$ by $(0,U,0)$ and $(0, 0, U)$ in turn The three equations of the type (5) must be identities in U, u, v, w

The solution of (5) (see Note 2) is, B being an arbitrary positive constant,
$$f(u,v,w)\,dudvdw = \frac{B\,dudvdw}{c^3\left(1-\dfrac{u^2+v^2+w^2}{c^2}\right)^2} \qquad (6)$$

90. This velocity-distribution has several remarkable properties. In the experience of any observer situated on an arbitrary member of the system, the system is a continuous one, containing a continuum of velocities defined by $|\mathbf{V}| = (u^2+v^2+w^2)^{\frac{1}{2}} < c$ The density of particles (in velocity-space) increases as $|\mathbf{V}|$ increases, tending to infinity as $|\mathbf{V}| \to c$, and the total number of particles $0 \leqslant |\mathbf{V}| < c$ is infinite The system contains no preferential velocity standard which could be called rest, all velocity-frames are equivalent, and the system possesses no velocity-centroid Whatever frame of reference is adopted, particles exist moving in every direction with all speeds up to that of light Solution (6) is in fact simply the expression in symbols of the non-existence of a preferential velocity-frame Any asymmetry in velocity-distribution would have picked out a

preferential direction, and had the number of particles been finite, a velocity-centroid could have been determined

Limiting spatial distribution

91. Now consider the spatial distribution of the system If t is the epoch measured from the moment when the system is first given, in the experience of an observer A, r the radial distance of any particle at epoch t, then for t large we have seen that

$$V \sim \frac{r}{t}, \tag{7}$$

where $V = |\mathbf{V}|$ In terms of coordinates x, y, z measured from A, this may be written

$$u \sim \frac{x}{t}, \qquad v \sim \frac{y}{t}, \qquad w \sim \frac{z}{t} \tag{8}$$

In the experience of this observer, the system thus tends to assume the spatial distribution

$$\frac{B\dfrac{dx}{t}\dfrac{dy}{t}\dfrac{dz}{t}}{c^3\left(1-\dfrac{x^2+y^2+z^2}{t^2}\right)^2},$$

or if n is the spatial particle-density for t large,

$$n\,dxdydz \sim \frac{Bt\,dxdydz}{c^3\left(t^2-\dfrac{x^2+y^2+z^2}{c^2}\right)^2}. \tag{9}$$

This formula expresses the fact that the swifter moving particles are found towards the confines of the system

Idealized system

92. Formula (9) invites consideration of the idealized system of particles of density-distribution

$$n\,dxdydz = \frac{Bt\,dxdydz}{c^3\left(t^2-\dfrac{x^2+y^2+z^2}{c^2}\right)^2}, \tag{10}$$

with the associated velocity-law

$$u = \frac{x}{t}, \qquad v = \frac{y}{t}, \qquad w = \frac{z}{t}, \tag{11}$$

x, y, z, t being coordinates and epoch assigned by the observer A on the particle at the origin of coordinates

It is to be noticed that this distribution of velocity and density satisfies identically the hydrodynamical equation of continuity,

$$\frac{\partial n}{\partial t}+\frac{\partial}{\partial x}(nu)+\frac{\partial}{\partial y}(nv)+\frac{\partial}{\partial z}(nw) = 0$$

as is easily verified Thus conservation of particle-number is observed by the flow

Let us consider the description of this system in the experience of any other observer B situated on some other particle of the system

93 Let B be on the particle moving with velocity $(U, 0, 0)$ and let him assign coordinates x', y', z' and epoch t' to the event on a particle P to which A assigns x, y, z, and t Further let B assign the velocity components u', v', w' to P, A the components u, v, w Then

$$x' = \frac{x-Ut}{(1-U^2/c^2)^{\frac{1}{2}}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t-Ux/c^2}{(1-U^2/c^2)^{\frac{1}{2}}}, \tag{12}$$

$$u' = \frac{u-U}{(1-uU/c^2)}, \quad v' = \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \quad w' = \frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2} \tag{13}$$

Hence, since $\qquad x/t = u, \quad y/t = v, \quad z/t = w,$

we have $\qquad u' = \dfrac{x'}{t'}, \quad v' = \dfrac{y'}{t'}, \quad w' = \dfrac{z'}{t'} \qquad (14)$

Thus the system possesses the same velocity-law for B using B's coordinates as it does for A using A's coordinates

94 Consider now the description of the density-distribution by B

A counts the particles inside $dxdydz$ at the common epoch t in his experience A particle at (x, y, z, t) in A's experience is found at (x', y', z', t') in B's experience A particle at $(x+dx, y+dy, z+dz, t)$ in A's experience is found at $(x'+\Delta x', y'+\Delta y', z'+\Delta z', t'+\Delta t')$ in B's experience, where

$$\Delta x' = \frac{dx}{(1-U^2/c^2)^{\frac{1}{2}}}, \quad \Delta y' = dy, \quad \Delta z' = dz, \quad \Delta t' = \frac{-U\,dx/c^2}{(1-U^2/c^2)^{\frac{1}{2}}}. \tag{15}$$

This same particle, at the assigned epoch t' in B's experience, will have had coordinates $(x'+dx', y'+dy', z'+dz')$, where

$$x'+dx' = x'+\Delta x'-u'\Delta t',$$
$$y'+dy' = y'+\Delta y'-v'\Delta t',$$
$$z'+dz' = z'+\Delta z'-w'\Delta t'$$

For during the interval $(t', t'+\Delta t')$ it has been moving with velocity (u', v', w') Hence, substituting from (14) and (15),

$$dx' = \frac{dx}{(1-U^2/c^2)^{\frac{1}{2}}}\left[1+\frac{x'}{t'}\frac{U}{c^2}\right],$$

$$dy' = dy+\frac{y'}{t'}\frac{(U/c^2)}{(1-U^2/c^2)^{\frac{1}{2}}}\,dx, \tag{16'}$$

$$dz' = dz+\frac{z'}{t'}\frac{(U/c^2)}{(1-U^2/c^2)^{\frac{1}{2}}}\,dx \tag{16''}$$

The first of these, on using (12), gives

$$dx' = dx(t/t') \tag{16}$$

Let now n' be the particle-density at (x', y', z') at epoch t' in B's experience Then for small volumes surrounding the particle (x, y, z) or (x', y', z'), since the particles counted by A at t will all be counted by B at t', we have if do' and do are elements of volume,

$$\int n'\,do' = \int n\,do$$

Consider now dx, dy, dz and dx', dy', dz' as local sets of coordinates, used for defining positions inside do and do' Then

$$\frac{do}{do'} = \frac{\partial(dx, dy, dz)}{\partial(dx', dy', dz')} = 1\Big/\frac{\partial(dx', dy', dz')}{\partial(dx, dy, dz)} = \frac{t'}{t}, \tag{17}$$

on differentiation of (16), (16′) and (16″) It follows by taking small volumes that

$$n' = n(t'/t)$$

But by (12)
$$t^2-\frac{x^2+y^2+z^2}{c^2} = t'^2-\frac{x'^2+y'^2+z'^2}{c^2} \tag{18}$$

Hence

$$n' = \frac{Bt}{c^3\left(t^2-\dfrac{x^2+y^2+z^2}{c^2}\right)^2}\frac{t'}{t}$$

$$= \frac{Bt'}{c^3\left(t'^2-\dfrac{x'^2+y'^2+z'^2}{c^2}\right)^2}$$

Thus B describes the density-distribution as

$$n'\,dx'dy'dz' = \frac{Bt'dx'dy'dz'}{c^3\left(t'^2-\dfrac{x'^2+y'^2+z'^2}{c^2}\right)^2} \tag{19}$$

This is identical in form with (10), and accordingly B's description of the density-distribution coincides with A's

95. Now A and B are equivalent observers, $A \equiv B$. We have shown that A's and B's descriptions of the velocity- and density-distributions of the system defined by (9) and (10) are identical. Thus $A \equiv B$. Further, A and B are any two particle-observers of the system. It follows that the system defined by (9) and (10) satisfies Einstein's cosmological principle, and that in the stricter sense that all particles occurring are equivalent and possess identical world-views.

This achieves a solution of the problem of constructing a system satisfying the cosmological principle in flat space. The calculation we have just performed can equally well be carried out using four-dimensional space-time, but this conceals the observational meaning of the calculations. The kinematical properties of the system will be considered shortly.

Accelerations

96. We have now to discuss the important question of whether the system of particles in motion which we have constructed can go on of itself, without constraints. So far the particles have been supposed compelled, by external means, to move with uniform relative velocities. Will they continue to do so if the constraints are removed?

97. In the presence of any distribution of particles, any free particle projected from a given point at a given epoch with a given velocity will move in general in some determinate way. For the only circumstances of 'release' of the particle which we can control are the epoch, position, and velocity of projection, all reckoned in the experience of an assigned observer.

Now suppose that P is one of our constrained particles and that a test-particle P' is released from P at some epoch with the velocity of P, in the reckoning of an observer O. Then P' originally coincides with P, and is at rest relative to P. Now describe the motion of P' from the point of view of an observer at P. The system described by (10) and (11) above is spherically symmetrical about the origin, i.e. about O. But P is equivalent to O. Hence the system is spherically symmetrical about P. Hence, viewed from P, there is no preferential direction. Hence the particle P' cannot separate from P, for to do so would be to select a preferential direction, and no such selection is possible. If P' had been projected from P with some definite velocity, the act of projection would have selected a direction. But P' has been projected from P, or released from P, with zero velocity relative

to P Hence P' remains in free motion permanently coinciding with P But P is in uniform motion relative to O Hence the free particle P' is in uniform motion relative to O, and therefore undergoes zero acceleration relative to O

But it is now impossible to distinguish between the motions of P and P' Hence the constraint compelling P to move uniformly may be removed, and P will follow the path of the free particle P' That is, P will continue to move with uniform velocity and zero acceleration in the experience of O

Having removed the constraint acting on P, we may pass to every particle of the system in turn, and remove its constraint without affecting its motion † The result is a system of particles moving *freely* with uniform relative velocities

The point of the argument is that P' and P are central in their own frames of reference If we suppose that in some way every other particle of the system 'acts' on P, then the whole 'action' must reduce to zero Therefore, since P is unaccelerated in its own frame, and since its own frame is in uniform relative motion with respect to O, P must be unaccelerated with respect to O Thus every particle of the system is necessarily unaccelerated But the assumption that the particles may 'act' on one another is essentially irrelevant The vanishing of the acceleration of each particle P is a consequence of the non-existence of a preferential direction in the frame of reference in which P is momentarily stationary

The acceleration-formula

98 I shall now arrive at the same conclusion by a different line of reasoning, by actually calculating the acceleration $d\mathbf{V}/dt$ of a free test-particle projected with given circumstances of projection The resulting formula is of fundamental importance in later developments It is derived by reasoning similar to that used in the earlier sections of Chapter IV

99 Let a free test-particle P' be projected from a particle P of the constrained system at epoch t with velocity $\mathbf{V}$ in the experience of an observer O situated on some particle of the system The acceleration $d\mathbf{V}/dt$ undergone by P' at the epoch of projection, as reckoned by O, will be denoted by $\mathbf{g}$, a vector O can observe particles freely pro-

† Alternatively, we may remove all the constraints simultaneously

jected from all particles P at all epochs t with all velocities $\mathbf{V}$ and determine their accelerations These accelerations can be enumerated by O as a function $\mathfrak{g}(\mathbf{P}, \mathbf{V}, t)$ of the vectors $\mathbf{P}$ and $\mathbf{V}$ and scalar t describing the circumstances of projection Thus we have

$$\frac{d\mathbf{V}}{dt} = \mathfrak{g}(\mathbf{P}, \mathbf{V}, t)$$

Now consider any other observer O' situated on some other particle of the system, and therefore in uniform relative motion with respect to O O' will describe the circumstances of projection of the test-particle as $\mathbf{P}', \mathbf{V}', t'$ If O' has velocity $(U, 0, 0)$ with respect to O, and if $\mathbf{V}$ is (u, v, w) and $\mathbf{V}'$ is (u', v', w'), then the acceleration $\mathfrak{g}'$ or $d\mathbf{V}'/dt'$ is given by differentiation of formulae (4) with respect to t' We have in fact if (f, g, h) are the components of the vector $\mathfrak{g}$, (f', g', h') the components of $\mathfrak{g}'$,

$$f = \frac{du}{dt}, \qquad g = \frac{dv}{dt}, \qquad h = \frac{dw}{dt}, \tag{20}$$

$$f' = \frac{du'}{dt'}, \qquad g' = \frac{dv'}{dt'}, \qquad h' = \frac{dw'}{dt'}, \tag{21}$$

whence

$$\left.\begin{aligned} f' &= \frac{(1-U^2/c^2)^{\frac{3}{2}}}{(1-uU/c^2)^3}f, \\ g' &= \frac{1-U^2/c^2}{(1-uU/c^2)^2}g+\frac{1-U^2/c^2}{(1-uU/c^2)^3}\frac{vU}{c^2}f, \\ h' &= \frac{1-U^2/c^2}{(1-uU/c^2)^2}h+\frac{1-U^2/c^2}{(1-uU/c^2)^3}\frac{wU}{c^2}f. \end{aligned}\right\} \tag{22}$$

From these formulae, then, from a knowledge of $\mathfrak{g} \equiv \mathfrak{g}(\mathbf{P}, \mathbf{V}, t)$, O' can determine the acceleration $\mathfrak{g}'$ as it would be measured by himself as a function of $\mathbf{P}, \mathbf{V}, t$ He can then express this result in terms of $\mathbf{P}', \mathbf{V}', t'$, say $\mathfrak{g}' \equiv \mathfrak{g}'(\mathbf{P}', \mathbf{V}', t')$ But O and O' are equivalent particle-observers, and therefore the acceleration $\mathfrak{g}'$ considered as a descriptive function of the variables $\mathbf{P}', \mathbf{V}', t'$ must be identical in form with the acceleration $\mathfrak{g}$ considered as a descriptive function of the variables $\mathbf{P}, \mathbf{V}, t$, i e

$$\mathfrak{g}' \equiv \mathfrak{g} \tag{23}$$

100. The physical meaning of this may be examined for a moment O records the acceleration of the particle projected from $\mathbf{P}$ with velocity $\mathbf{V}$ at epoch t as $\mathfrak{g}(\mathbf{P}, \mathbf{V}, t)$ O' records the acceleration of the same particle, in his coordinates, as $\mathfrak{g}'(\mathbf{P}', \mathbf{V}', t')$ But the latter

acceleration must be equal to that of a free particle projected from a particle P' with velocity $\mathbf{V}'$ at epoch t' *in the experience of O* For O and O' are equivalent We are here using the principle of relativity in its primitive meaning if O and O' are equivalent and observe identical world-pictures, then the law expressing the acceleration of a freely moving particle in A's measures must be identical *in form* with

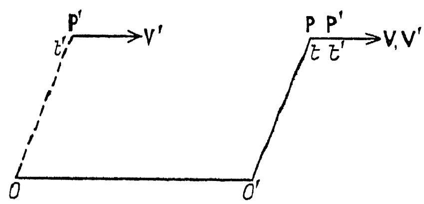


FIG 10 Enumeration of accelerations of free particles by two observers O, O' such that $O \equiv O'$

the law expressing the acceleration of a freely moving particle in O''s measures The acceleration-law must be unaltered *in form* when we transform from one observer to another equivalent observer †

101. The scalar numbers f, g, h are functions of the 7 variables x, y, z, u, v, w, t, which we write for brevity in the form

$$f \equiv f(\mathbf{P}, \mathbf{V}, t), \qquad g \equiv g(\mathbf{P}, \mathbf{V}, t), \qquad h \equiv h(\mathbf{P}, \mathbf{V}, t)$$

The identity (23), together with the transformation formulae (22), now gives

$$f(\mathbf{P}', \mathbf{V}', t') = \frac{(1-U^2/c^2)^{\frac{3}{2}}}{(1-uU/c^2)^3} f(\mathbf{P}, \mathbf{V}, t),$$

$$g(\mathbf{P}', \mathbf{V}', t') = \frac{1-U^2/c^2}{(1-uU/c^2)^2} g(\mathbf{P}, \mathbf{V}, t) + \frac{1-U^2/c^2}{(1-uU/c^2)^3} \frac{vU}{c^2} f(\mathbf{P}, \mathbf{V}, t),$$

$$h(\mathbf{P}', \mathbf{V}', t') = \frac{1-U^2/c^2}{(1-uU/c^2)^2} h(\mathbf{P}, \mathbf{V}, t) + \frac{1-U^2/c^2}{(1-uU/c^2)^3} \frac{wU}{c^2} f(\mathbf{P}, \mathbf{V}, t) \qquad (24)$$

† The tensor formulation of 'laws of nature' in general relativity is an expression of conservation of *quantity* when the value of a tensor is zero in one system of coordinates, it is so in all On the other hand our use of the principle of relativity is one of conservation of *functional form* In this respect our application is more intimately connected with the foundations of relativity. For relativity starts by establishing the conservation, under certain circumstances, of the *form* of the quadratic differential expression $c^2dt^2-(dx^2+dy^2+dz^2)$ General relativity proceeds to generalize this result by concentrating on the conservation of *value* of this expression, calling it ds^2 But in doing so it loses sight of the remarkable property that not merely two *numerically* equal quadratic differential expressions exist connecting the descriptions of the same pair of neighbouring events by two observers, but the two expressions are identical in *form* In Chapter II we have seen how to generalize this result for relatively accelerated 'equivalent' observers

These constitute three functional equations for the unknown functions f, g, h We have six other functional equations obtained by replacing $(U, 0, 0)$ by $(0, U, 0)$ and $(0, 0, U)$

The solution of these functional equations is found in Note 3 Let X, Y, Z, ξ be four functions of x, y, z, t, u, v, w defined as follows

$$X = t^2 - \frac{\mathbf{P}^2}{c^2}, \qquad Y = 1 - \frac{\mathbf{V}^2}{c^2}, \qquad Z = t - \frac{\mathbf{P}\ \mathbf{V}}{c^2} \tag{25}$$

or

$$X = t^2 - \frac{x^2+y^2+z^2}{c^2}, \qquad Y = 1 - \frac{u^2+v^2+w^2}{c^2}, \qquad Z = t - \frac{ux+vy+wz}{c^2}, \tag{25'}$$

and

$$\xi = Z^2/XY \tag{26}$$

Under the transformation from an observer O to an observer O' moving with relative velocity $(U, 0, 0)$, the functions X, Y, Z for the same motion-event (P, V, t) or (P', V', t') obey the transformation laws

$$X' = X, \tag{27}$$

$$Y' = Y\frac{1-U^2/c^2}{(1-uU/c^2)^2}, \tag{27'}$$

$$Z' = Z\frac{(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \tag{27''}$$

so that X, Z^2/Y, and $\xi = Z^2/XY$ are invariants under the transformation Then the solution of the set of functional equations is

$$\mathbf{g} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi), \tag{28}$$

where $G(\xi)$ is a function of ξ not yet determined

Now for any freely moving particle, the motion at any point of its path is the same as if it had been freely projected at that point, at the same epoch, with the velocity with which it is passing through that point The equations determining the motion of a free particle are accordingly

$$\frac{d\mathbf{P}}{dt} = \mathbf{V}, \qquad \frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi) \tag{29}$$

102. If in (28) we take $\mathbf{V} = \mathbf{P}/t$, we have $\mathbf{g} = 0$. Hence a free particle projected with the velocity $\mathbf{V} = \mathbf{P}/t$ is unaccelerated The solution of the equations $d\mathbf{V}/dt = 0$, $d\mathbf{P}/dt = \mathbf{V}$ is precisely $\mathbf{V} =$ constant, $\mathbf{P} = \mathbf{V}t$ Accordingly a test-particle released from one of the fundamental or given particles $\mathbf{V} = \mathbf{P}/t$ at rest relative to it remains at rest

relative to it There is therefore no distinction between the motion of the given, constrained particle and the motion of a free particle coincident with it Coincidence persists. Hence the constraint acting on the given particle may be removed without affecting its motion The constraints on all the given particles may be released in turn, and we are left with a natural, unconstrained motion Thus the system of particles moving according to (11) and distributed according to (10), in the experience of the particle-observer at the origin, maintains this motion and distribution, and the motion and distribution are described in the same way by any other particle-observer of the system For any other observer they follow the same laws †

103. Whether the particles 'act' on one another in any way, i e whether a 'law of gravitation' exists, is an irrelevant question We may, if we like, assume that they do so act on one another But provided the 'law of gravitation' is relativistic, i e provided it is described in identical terms by equivalent observers ($A \equiv B$) who describe the system equivalently ($A \equiv B$), the free motion is precisely as we have stated it The particles could only move otherwise under a non-relativistic 'law of gravitation' The motion we have obtained is compatible with *any* relativistic formulation of the 'law of gravitation', if such be supposed to exist We may impose any kind of interaction between the particles we like, involving any number of 'universal' constants, $\gamma, \lambda, \ldots$, we may adopt any 'theory' of gravitation we choose, but provided the interaction is relativistic, provided the

† An exactly similar procedure is tacitly followed in treatments based on general relativity A gravitational field is contemplated which originates from a certain distribution of matter in motion The geodesics, or paths of free particles in this field, are then calculated Lastly, particular cases of these geodesics are shown to be followed by the particles originally given to be present to 'cause' the field, which previously had prescribed motions, and thus *a priori* required to be constrained in order to follow these motions It then follows that no constraints are necessary, so that the original constraints, if any, may be removed (I am referring to the case of zero pressure) The whole point, in both the general relativity treatments and in my kinematic treatment, is that the prescribed motions of the matter 'producing the field' are particular cases of the paths of free particles in the field It is true that in the general relativity treatments, the 'field' or metric is often considered merely kinematically, without specification of the matter present producing it, this matter being later determined from the 'field' equations Since the field equations are indeterminate to the extent of an unknown constant λ, the amount of matter present is also indeterminate But in any case its motion follows the geodesics defined by taking constant values of the chosen coordinates It is really illogical to associate a gravitational field with the presence of matter, but at the same time to begin an investigation with consideration of an abstract field or metric, defining a set of geodesics without first saying how much matter is present and what it is doing

theory is compatible with its description by equivalent observers in equivalent terms, the motion is unaffected † None of our arguments are affected in any way We have thus constructed what may be called a gravitating system of particles, in the flat space and Newtonian time of any occurring particle-observer, whose motion is the same on any theory of gravitation Specialization of the 'law of gravitation' might affect the function $G(\xi)$ (*a priori*) and so affect the general paths of free test-particles, but will not affect the motion of the given or fundamental particles given to be present and supposed to originate the field ‡

This is a most satisfactory result An essential phenomenon like gravitation can only be said to be understood when it has been shown to follow inevitably, without arbitrary assumptions or arbitrary constants, from the compatibility of the observations which the different particle-observers occurring in it can make on one another The Newtonian formulation of gravitation is a specific assumption of the form of the law which accounts for a certain class of observed motions Einstein's formulation of the *form* of the law of gravitation is a possible restriction on the motions (or alternatively on the possible density-distributions compatible with the prescribed motions) containing an arbitrary constant λ We have not obtained a general formulation of the law of gravitation, but we have obtained a definite system of particles in motion whose subsequent motions are independent of the *form* of the 'law of gravitation', if such is supposed to exist, and so independent of any assumption about the existence or non-existence of the 'cosmical constant' λ We have not 'left out' gravitation The uniformity of the motions in our system can be interpreted if we please as a balance, at each particle occurring, of the 'forces' 'due to' every other particle occurring The manner in which this resultant zero is analysed into 'action at a distance' is arbitrary, and the

† See Additional Note, p 112

‡ This may be contrasted with general relativity There, given the metric, the geodesics or free paths are prescribed independently of the field equations adopted Thus in general relativity, when we begin with a metric, the free paths are known, but the matter 'causing' them depends on the particular field equations adopted, i e on the law of gravitation adopted, in our treatment, the matter present, and its motion, are known and definite, but the particular forms of the trajectories of *additional* test particles *may* depend on the law of gravitation adopted, as far as we have yet gone This brings into evidence the fundamental differences between our kinematic treatment and general relativity We shall later actually determine the function $G(\xi)$ for statistical systems satisfying the cosmological principle in terms of the matter given to be present

analysis could be carried out in different ways But no imposition of any relativistic law of gravitation can disturb the set of motions we are discussing

104. But we have gone farther than merely to determine the motions occurring We have actually found a considerable restriction on the acceleration of *any* free test-particle The acceleration function $\mathbf{g}(\mathbf{P}, \mathbf{V}, t)$ can be taken to be a descriptive formulation of the 'law of gravitation' for the 'effect' of all the given particles on a free test-particle It can be taken as summarizing the resultant action on the free particle of all the given particles present This function we have shown to be of the form (28), reducing $\mathbf{g}$ from a descriptive function of 7 variables to a single undetermined function of a single variable ξ, namely $G(\xi)$ Just as the motion of the prescribed or given particles is independent of any specific formulation of the 'law of gravitation', or 'law of interaction', so the motion of a free particle is governed by an acceleration-law of the form (28) independent of any specific theory of gravitation As far as we have at present gone, the form of the residual factor $G(\xi)$ might depend on the precise 'law of gravitation' adopted It will in any case be expected to depend upon the arbitrary multiplying constant B occurring in the density-distribution-law (10) We shall later show that the equations of motion (29) can be integrated in full without any knowledge of the form of $G(\xi)$, in a form involving explicitly the necessary six constants of integration corresponding to arbitrary values of the position $\mathbf{P}$ and velocity $\mathbf{V}$ at an arbitrary epoch of projection t, and we shall find it possible to state a large number of properties of the trajectories so found without further particularization of the function G We shall thus have gone a long way towards describing the complete 'gravitational field' of the system, not only in regard to the accelerations of the *given* particles occurring (which we know completely, namely they are zero) but in regard to the accelerations of free test-particles used to explore the field Formula (28) *is* the form of the law of gravitation as explored by free test-particles at large in the given system It rests on no assumption as to the existence of action at a distance or otherwise, it rests on no causative principle, it is not derived from any assumption of an 'effect' on 'space-curvature' or geometry due to the presence of matter, it involves no introduction of an ether Whatever law of gravitation is imposed, adopted, or formulated, equivalent

observers O must describe the function $\mathbf{g}$ in the same terms, i e must arrive at a single universal function $\mathbf{g}$, and it is then a consequence of the equivalence of the observers that $\mathbf{g}$ is of the form (28) A general 'law of gravitation' can be none other than a compact statement of all the motions of free test-particles released in the presence of a distribution of given particles, including the given particles considered as free particles We have effectively formulated the law of gravitation for a certain system, namely that defined by (10) and (11), in the form of a compact statement of a restriction on the motion of all possible free test-particles

Physical interpretation of the acceleration-formula

105. Let us examine (28) more closely The first thing which will strike the reader is the dependence of the acceleration on velocity $\mathbf{V}$, which is

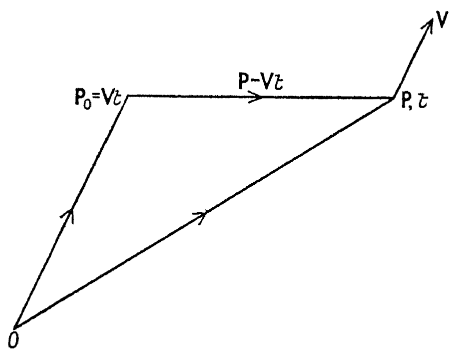


FIG 11 Physical interpretation of the acceleration-formula

apparently foreign to all classical ideas as to gravitation We proceed to interpret physically the occurrence of $\mathbf{V}$ in the factor $\mathbf{P}-\mathbf{V}t$

Consider an observer O on a given particle taken as origin Then a free test-particle moving through P with velocity $\mathbf{V}$ at time t, all in the experience of O, possesses an acceleration $d\mathbf{V}/dt$ which by (28) is parallel (or antiparallel) to the vector $\mathbf{P}-\mathbf{V}t$ It is therefore directed towards (or away from) the particle P_0 of position-vector $\mathbf{V}t$ with respect to O Thus in Fig 11, the acceleration of the $\mathbf{V}$-particle at P at time t is along the vector $P_0 P = \mathbf{P}-\mathbf{V}t$

But P_0 is the position, at time t, of that one of the given particles which has been moving with a constant velocity equal to $\mathbf{V}$ To the particle-observer P_0, the $(\mathbf{P}, \mathbf{V}, t)$ projected particle is momentarily at rest, for the projected particle (travelling with $\mathbf{V}$) and P_0 possess the

same velocity in O's reckoning, and so in the reckoning of any equivalent particle, in particular in the reckoning of P_0, to whom both P_0 and the free test-particle under consideration are at rest Now O is the centre of spherical symmetry of the whole system to the particle-observer at O Hence P_0, an equivalent particle, is the centre of spherical symmetry of the whole system to the particle-observer at P_0, and so the centre of spherical symmetry of the whole system to any particle-observer at rest relative to P_0 It follows that the *acceleration of the free particle in flight at P with velocity V is directed towards (or away from) the centre of symmetry of the whole system in that frame in which the free particle is momentarily at rest*

106. This again is a most satisfactory result A particle-observer at P moving with $\mathbf{V}$ ($\mathbf{V} \neq \mathbf{P}/t$) can survey the system regarding himself as momentarily at rest, and ascertain its centre of symmetry He thus picks out P_0 He then finds that his acceleration is along the line joining himself to P_0 Had we adopted some 'theory' of gravitation, we should inevitably have concluded on grounds of symmetry that the 'resultant action' of all the particles would be directed towards the centre of the system in the experience of O We have not made this assumption, but our conclusion is compatible with the results of this assumption Should $\mathbf{V}$ happen to be $\mathbf{P}/t$, P_0 coincides with P, P finds himself at the centre of symmetry of the system in the frame in which he is instantaneously at rest, and so would be led to say that the resultant acceleration, having no preferential direction, must be zero This is our earlier result

107 If $\mathbf{V}$ is not equal to $\mathbf{P}/t$, $d\mathbf{V}/dt$ is not zero, $\mathbf{V}$ alters with t, $\mathbf{V}t$ alters with t, and the position P_0 has a definite locus in the system in the experience of O P_0, the apparent centre of the system to the free particle P, does not permanently coincide with any *given* particle of the original system, but moves from particle to particle P, the free particle, may be said to follow a curve of pursuit of the apparent centre P_0 (if the acceleration has the sense of PP_0) P may be described as steadily falling towards the centre of the system, reckoned always in the frame in which P is momentarily at rest Whether P ever overtakes P_0 will be a matter for future investigation.

108. In the meantime we have obtained a physical interpretation of the occurrence of the vector-factor $\mathbf{P}-\mathbf{V}t$ in the acceleration-

formula for P This factor is an embodiment in mathematics of the circumstance that the system possesses always an apparent centre in any frame

According to (10), the given density-distribution for small $r = (x^2+y^2+z^2)^{\frac{1}{2}}$ is locally *homogeneous*, of particle density

$$\frac{Bt}{c^3t^4}\left(1+2\frac{r^2}{c^2t^2}+\quad\right)$$

$$\sim B/c^3t^3$$

But its density increases outwards in the experience of any of the equivalent observers O To O there appears a density-gradient near any distant particle-observer O', reckoned in the coordinates used by O, but in the same vicinity, reckoned in the coordinates used by O', the density-gradient of the same particles flattens out to zero, and the system to O' is again locally homogeneous near O'. The coordinates used, whether by O or O', coincide always with those habitually employed in physics We see, first, how far removed from *actual homogeneity in experience* is the density-distribution of a system which satisfies the cosmological principle, i e which possesses *homogeneous descriptions* by O and O' We see, secondly, that without any recourse on our part to anthropomorphic arguments, the free particle P has recognized this *de facto* non-homogeneity, and assumed an acceleration directed towards the centre of spherical symmetry of this recognizably non-homogeneous system

Meaning of homogeneity

109. The cosmic time t' of an event t occurring on a fundamental particle at P moving with velocity $\mathbf{V} = \mathbf{P}/t$ is by our earlier results given by

$$t' = \left(1-\frac{\mathbf{V}^2}{c^2}\right)^{\frac{1}{2}}t = \left(1-\frac{\mathbf{P}^2}{c^2t^2}\right)^{\frac{1}{2}}t.$$

The particle-density near O, at epoch t, is B/c^3t^3. At O' at O''s epoch t', in O''s reckoning, it is $B/c^3t'^3$ Thus O and O' assign the same numerical values to their local densities at equal local times, and thus the system is strictly homogeneous in particle-density in the sense employed in current relativistic cosmology. But this featureless homogeneity reckoned locally in local times gives no opportunity for a physical interpretation of the acceleration-formula, which rests essentially on the circumstance that neither O nor O', using their own coordinates for distant as well as near events, will *see* the system as

homogeneous This shows the marked advantage of employing the coordinates we have defined and adopted, namely the ordinary coordinates of physics Current relativistic cosmology can equally well determine the acceleration of a free particle at P, and ascertain its direction But in relativistic cosmology the system is described as absolutely homogeneous, and therefore centreless, and so no physical meaning can be immediately given to the acceleration-direction An observer on one of the fundamental particles, using the ordinary coordinates and procedure of ordinary physics, will, however, find the system to be distributed with spherical symmetry round himself, with himself as centre, and any other observer, moving with any velocity whatever at any point at any time, will see himself as excentric to the system, and find his acceleration directed towards (or away from) the apparent centre in his own view

The fundamental observer at O reckons the particle-density at P at time t as

$$\frac{Bt}{c^3\left(t^2-\frac{r^2}{c^2}\right)^2}$$

The fundamental observer at P reckons his local time as

$$t' = \left(1-\frac{r^2}{c^2t^2}\right)^{\frac{1}{2}}t$$

and so estimates the density near himself as $B/c^3t'^3$ or

$$\frac{B}{c^3}\frac{1}{\left(t^2-\frac{r^2}{c^2}\right)^{\frac{3}{2}}}.$$

The ratio is given by

$$\frac{\text{density at } P \text{ in } P\text{'s reckoning}}{\text{density at } P \text{ in } O\text{'s reckoning}} = \left(1-\frac{r^2}{c^2t^2}\right)^{\frac{1}{2}} = \left(1-\frac{V^2}{c^2}\right)^{\frac{1}{2}}$$

and thus P reckons the density as smaller than O does, at the same events at P This arises because to O the distribution is subject to the Lorentz-contraction due to the velocity V. The density at P at a given event at P is smaller in the reckoning of the fundamental observer at P than in the reckoning of any other fundamental observer

Local accelerations

110. Given any free particle, we can now always find a fundamental particle-observer relative to whom the free particle is momentarily at rest Consider then a free particle momentarily at rest with regard

to a fundamental particle O. Let its position-vector $\mathbf{P}$ or $\mathbf{r}$ with respect to O have modulus r. Then its acceleration is, putting $\mathbf{V} = 0$ in (28),

$$\mathbf{g} = \frac{d\mathbf{V}}{dt} = \frac{\mathbf{r}}{t^2-\frac{r^2}{c^2}} G\left(\frac{t^2}{t^2-\frac{r^2}{c^2}}\right). \tag{30}$$

For r small, this reduces approximately to

$$\frac{d\mathbf{V}}{dt} \sim \frac{\mathbf{r}}{t^2} G(1) \tag{31}$$

This is of the same form as the Newtonian acceleration at the same point. For the local particle-density is B/c^3t^3, and if m is the mass of a particle, the Newtonian acceleration at distance r is

$$-\frac{\gamma}{r^2}\tfrac{4}{3}\pi r^3 \frac{mB}{c^3t^3},$$

and directed inwards, where γ is the Newtonian constant of gravitation. This is simply

$$-\frac{4}{3}\frac{\pi\gamma mBr}{c^3t^3}, \tag{32}$$

which agrees with (31) provided

$$G(1) = -\frac{4}{3}\frac{\pi mB}{c^3}\frac{\gamma}{t} \tag{33}$$

Now $G(1), m, B$ are constants. Hence the total effect of all the material present is reproduced as regards local accelerations provided $\gamma \propto t$. The Newtonian constant γ is of course defined to be independent of t. Our evaluation of an effective γ, judged not from two ideal particles alone in space (as in the Newtonian definition) but from the actual local acceleration of a test-particle in the presence of a certain distribution of particles in motion, gives γ increasing secularly as the epoch measured from the natural zero of time. The comparison suggests that $G(1)$ is a negative number.†

111. The density-distribution given by (10) with the motion (11) extends throughout the expanding sphere of radius $r = ct$. If we extrapolate the observed local density B/c^3t^3 so as to construct a distribution strictly homogeneous in the experience of the particular observer O concerned, and filling the sphere of radius $r = ct$, the

† The local density (at the observer), namely mB/c^3t^3 in mass-units, is by (33) equal to $-G(1)/(\frac{4}{3}\pi\gamma t^2)$ or $[-G(1)]\times 10^{-27}$ gramme cm.$^{-3}$ This is discussed in Chapter VI.

total number of particles in the extrapolated distribution will be

$$\tfrac{4}{3}\pi(ct)^3 \times \frac{B}{c^3t^3}$$

$$= \tfrac{4}{3}\pi B, \tag{34}$$

which is a constant, independent of t. Thus the 'extrapolated homogeneous system' has a constant particle population. Its mass is

$$\tfrac{4}{3}\pi Bm$$

$$= -\frac{G(1)c^3t}{\gamma}. \tag{35}$$

For $t = 0{\cdot}6 \times 10^{17}$ seconds, taking $c = 3 \times 10^{10}$ and $\gamma = 6{\cdot}66 \times 10^{-8}$ in customary units,† the mass of the extrapolated homogeneous system is

$$-G(1)\frac{0{\cdot}6 \times 10^{17} \times 27 \times 10^{30}}{6{\cdot}66 \times 10^{-8}}$$

$$= 2{\cdot}43 \times 10^{55}[-G(1)] \text{ grammes.}$$

The value $2{\cdot}4 \times 10^{55}$ grammes is that usually assigned to the mass of the universe on those cosmological theories which adopt a curved finite space in which to represent the universe. Its analytical expression is simply c^3t/γ, as here. Eddington has constructed a considerable theoretical edifice on the basis of this number $2{\cdot}4 \times 10^{55}$ grammes, the mass of $1{\cdot}5 \times 10^{79}$ protons, but the number appears to be devoid of any profound theoretical significance. It is merely the mass which would be assigned to the system of the universe by an observer who measured the local density, near himself, ascertained t, the distance of the present epoch from the natural zero of time, by means of the velocity-distance proportionality, and then erroneously supposed the system to be homogeneous in the ordinary sense of physics. The mass-dimensional number c^3t/γ, where t is the epoch of observation, is common to all evaluations of the mass that can be *observed* at epoch t, on all relativistic cosmologies, as we shall see. But this does not justify us in assigning this as the mass of the universe.

On the kinematic model here outlined the *total* particle-population of the system is infinite. This is easily seen by integrating n as given by (10) over the interior of the sphere of radius $r = ct$. The 'locally-extrapolated' homogeneous particle-population is, as we have seen, finite and constant. On the relativistic theories, on the other hand, c^3t/γ increases with epoch of observation t. Certain of these theories

† This assumes that we are justified in taking for the intergalactic value of γ the value derived by observations inside the solar system.

imply a steady increase with t in the total mass observable at epoch t, which indeed $\to \infty$ as $t \to \infty$. This will be established later.

Properties of the 'hydrodynamic' or simple kinematic system

112. The 'hydrodynamic' system of flow defined by

$$\mathbf{V} = \mathbf{P}/t, \qquad n\,dxdydz = \frac{Bt\,dxdydz}{c^3(t^2-r^2/c^2)^2} \tag{36}$$

has many close relations with the world-systems proposed in current 'relativistic' cosmology other than that just indicated. These can only be discussed after the 'relativistic' systems have been described. It is convenient now to summarize the properties of the system, some of which we have already discussed.

(1) The system is described in the same way by the same formulae (36) by any observer situated on any particle of the system, using his own coordinates, in flat space constructed out of his own clock-measures.

(2) The system is spherically symmetrical round any particle of the system, in the experience of the observer attached to that particle.

(3) The particle-density is locally homogeneous near any given particle-observer O of the system, in O's reckoning. Departures from homogeneity are of the second order in r/ct.

(4) The particle-density, in the reckoning of any particle-observer O, at any given epoch t, increases outwards.

(5) Near O, at any fixed distance, the particle-density decreases at a rate inversely proportional to the cube of the time.

(6) The system is contained at any epoch t within a finite expanding sphere centred round any particle-observer O, of radius $r = ct$, where t is the age of the system in O's reckoning. The radius of this sphere increases with the speed of light.

(7) As the distance r tends to ct, i.e. for points nearer and nearer the expanding light-sphere, the particle-density tends to infinity.

(8) The total number of particles in the system is infinite.

(9) The members of the system form at any epoch t in the experience of any particle-observer O an *open* set of points of which every point of the expanding sphere $r = ct$ is a limiting point. Every particle of the system is completely surrounded by other particles. No particle stands on the 'edge' of the system.

(10) Every particle of the system is in uniform radial motion outward from any arbitrary particle O of the system, and the acceleration

of every particle of the system is zero But the acceleration of a freely projected particle, other than the given particles, is not zero

(11) The domain occupied by the system, though finite in volume, has all the properties of infinite space, since its boundary is for all time entirely inaccessible by any hypothetical observer travelling with a speed not exceeding the speed of light

(12) The velocities of different particles at any one epoch are proportional to the distances of the particles from any assigned particle taken as origin, and tend to the velocity of light as the distance tends to ct

(13) If the particles are supposed to be luminous, then the luminosity near the expanding boundary approaches zero, since the particles are receding with nearly the speed of light (see Note 7)

(14) At an event E, at epoch t, on a particle P moving with velocity $\mathbf{V} = \mathbf{P}/t$ in the experience of a given particle-observer O, the local reckoning of the epoch (P's reckoning) is $t' = t(1-V^2/c^2)^{\frac{1}{2}}$ As $|\mathbf{V}| \to c, t' \to 0$ Thus if a particle is supposed to have an evolutionary history, the events reckoned by O to be now occurring on it are events earlier in the local time-reckoning of P.

(15) If t' is the local epoch at P of the event assigned epoch t by O, then O observes this event at epoch t_2 given by

$$t_2 = t+\frac{r}{c} = t\left(1+\frac{V}{c}\right) = t'\left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}} = st', \qquad (37)$$

where s is the Doppler shift at P observed by O.†

(16) At the event E at P, of epoch t in O's experience and of epoch t' in P's experience, the radius of the system is ct in O's experience and ct' in P's experience Thus the radius of the system for the same given event depends on the observer who observes that event It takes its smallest value for the observer close to the event in question. The system has accordingly no definite age or radius at any assigned event, the age t and radius ct depending on the epoch assigned to the event, which depends in turn on the observer making the assignment.

(17) A particle-observer O at the moment of experiencing an event E_2 at himself is at a much later stage of his own experience, reckoned in his own time-scale, than P is in his (P's) time-scale at the event E_1 at P which O is then observing

Nevertheless it is impossible for P to obtain foreknowledge of his own future by messages from O relating O's experiences For if the

† This is a particular case of the formula given in §§ 21, 22, Chapter II.

event E_2 at O has an epoch t_1 (in O's reckoning), the message will be received by P at an epoch t_2 in O's reckoning, where

$$Vt_2 = c(t_2-t_1)$$

or

$$t_2 = \frac{c}{c-V}t_1$$

Hence, on P's reckoning,

$$t_2' = (1-V^2/c^2)^{\frac{1}{2}}t_2 = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}}t_1 > t_1$$

Thus by the time (t_2') that P has received a communication as to O's experiences at O's time t_1, P has already had an experience extending in his own time-scale later than t_1 It is therefore impossible for P to attempt to foretell his own evolutionary future by communications from O

The world-map

113. Many of these properties are simply displayed in the diagram given in Plate I, which represents a cross-section of the system as mapped by an observer O at epoch t in his experience O is of course central, and maps from his observations a set of particles, filling the interior of the sphere $r = ct$ and more and more crowded towards its outer boundary

The world-picture

114. This map should be carefully distinguished from the world-picture which O would *photograph* at epoch t In the world-map, t is the common epoch of all the events mapped In the world-picture, t is the epoch of observation, and all the different events photographed have different epochs t_1 in O's reckoning The relation between the world-map and the world-picture is readily found as follows

The number of particles at epoch t inside a solid angle $d\omega$ and between distances r and $r+dr$ is

$$\frac{Btr^2\,drd\omega}{c^3(t^2-r^2/c^2)^2}$$

At epoch of observation t, these appear in the photograph as between distances r_1 and r_1+dr_1, where

$$t_1+r_1/c = t, \qquad V = \frac{r_1}{t_1} = \frac{r}{t}$$

or

$$r_1 = \frac{r}{1+r/ct}, \qquad r = \frac{r_1}{1-r_1/ct},$$

and

$$dr = \frac{dr_1}{(1-r_1/ct)^2}$$

Hence the number may be written

$$\frac{Bt\left(\dfrac{r_1}{1-r_1/ct}\right)^2 \dfrac{dr_1}{(1-r_1/ct)^2}}{c^3\left[t^2-\dfrac{r_1^2/c^2}{(1-r_1/ct)^2}\right]^2}\,d\omega$$

$$= \frac{Br_1^2\,dr_1\,d\omega}{c^3t(t-2r_1/c)^2} \tag{38}$$

This is a system confined within the expanding sphere of radius $r_1 = \frac{1}{2}ct$, and increasing towards its boundary The outward apparent

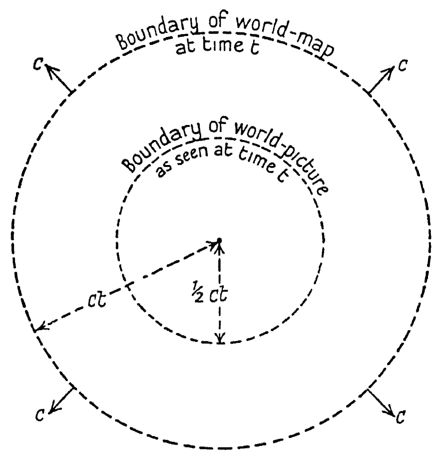


FIG 12 World map and world-picture for the simple kinematic system

increase of density near the observer O in the photograph is given by expanding the last expression as

$$\nu(r_1, t) \sim \frac{B}{c^3t^3}\left(1+\frac{4r_1}{ct}\right)$$

$$\sim n_0(1+4V/c), \tag{39}$$

in accordance with the general result of Chapter IV, equation (14).

It is clear without calculation that the radius of the system as observed at time t must be $\frac{1}{2}ct$ and not ct For the particles which are near the expanding boundary will be moving with nearly the speed of light, and so by the time their light-signals reach the camera they will have travelled as far as the light-signals have travelled, i e have doubled their distances, accordingly, since their distances in the world-map are ct, in the world-picture they must be $\frac{1}{2}ct$

The relations are shown in Fig 12

115. Still another aspect of the world-picture is worth calculating The velocity-distribution law

$$\frac{B\,du dv dw}{c^3(1-V^2/c^2)^2}$$

or its equivalent form
$$\frac{BV^2\,dV d\omega}{c^3(1-V^2/c^2)^2}$$

may appear very unlike those customarily met with in physics, such as Maxwell's law in the kinetic theory of gases But it is of a totally different character, since it expresses a velocity-count for particles whose position and velocity are correlated The best way of envisaging this velocity-distribution is to determine the number of particles possessing Doppler shifts within given limits If s is the Doppler shift ratio for a particle moving with velocity V, then from

$$s^2 = \frac{1+V/c}{1-V/c}, \qquad \frac{V}{c}=\frac{s^2-1}{s^2+1}$$

we have
$$\frac{dV}{c} = \frac{4s\,ds}{(s^2+1)^2}$$

Hence the number of particles is

$$B\left(\frac{s^2-1}{s^2+1}\right)^2 \frac{4s\,ds}{(s^2+1)^2}\left[\frac{(s^2+1)^2}{(s^2+1)^2-(s^2-1)^2}\right]^2 d\omega$$
$$= B\,d\omega\,\frac{(s^2-1)^2\,ds}{4s^3} \tag{40}$$

This formula is independent of any conventions as to coordinates, and relates solely the numbers counted for different ranges of Doppler effect It will be used in due course to compare the present distribution with other proposed distributions as regards observable phenomena

116 The advantage of the system discussed in this chapter is that it solves, in flat space and in a manner easily pictured, the problem of finding a set of particles in motion which are all equivalent to one another, and which therefore possess no accessible boundary Each member of the system is equally surrounded by other members receding from it The system, in the view of any given member, contains a boundary or rim or edge, and though the particles tend to invisibility as more and more distant ones are viewed, there is a maximum distance which is never exceeded at any given epoch t, but

which increases proportionally to t, the total radius occupied, in the view and reckoning of any arbitrary member, being finite (of volume $\frac{4}{3}\pi c^3t^3$) but expanding It is the system itself which is expanding To speak of the space itself as in a state of expansion is meaningless, for no meaning can be given to 'expanding space' or 'expanding emptiness' The system is expanding in the ordinary sense in which a rod expands when heated the distance between any two members is increasing. The distant particles, in the view of any given particle, are enormously crowded together, but if we journeyed towards the apparently crowded region we should find it not more crowded but less crowded than the region we had left If a messenger left our own neighbourhood at time t (reckoned by ourselves) with the speed of light, he would arrive at the distant particle P, of speed V and at present distance r from us, at time t_1 in our reckoning, where

$$t_1 = t+\frac{Vt_1}{c}$$

or

$$t_1 = \frac{t}{1-V/c},$$

where

$$V = r/t$$

The density at ourselves at the moment he left is B/c^3t^3 The density at his destination on his own reckoning, at the moment of his arrival, is $B/c^3t_1'^3$, where

$$t_1' = (1-V^2/c^2)^{\frac{1}{2}}t_1,$$

and thus is equal to

$$\frac{B}{c^3t^3}\frac{(1-V/c)^3}{(1-V^2/c^2)^{\frac{3}{2}}} = \frac{B}{c^3t^3}\left(\frac{1-V/c}{1+V/c}\right)^{\frac{3}{2}}$$

which is less than the density at ourselves, in our or his reckoning, at the moment he left. Thus it is never possible to experience, by journeying to a distant apparently more crowded region, a density greater than that experienced at home. Throughout the system's career, dilution of density is a one-way process, however the observer travels To experience a greater local density than that at home would be in effect, since all localities are equivalent, to re-experience the past, and this is for ever impossible We have already seen that neither can a distant observer receive messages about his own future evolutionary experiences, by messages from other observers who are in effect experiencing later experiences at the moment at which the message starts, before he has already experienced experiences later

than those about which the message tells. Thus neither re-experience of the past, nor foreknowledge of the future by direct message, are possible. The past is a closed book, the future an unreadable one—though the future march of density can be found by calculation. The Lorentz contractions and changes of time-reckoning throughout conspire to frustrate any effort to circumvent the one-directional character of the temporal experience, which indeed we posited as a basis of our calculations.

117. The system we have described is the extension to a system of material moving particles of that phenomenon which Einstein placed at the beginning of his investigations in relativity: the phenomenon that if two observers in uniform relative motion send out an expanding light wave, each will for ever see himself as the centre of the spherical wave, in his own reckoning. In our system each uniformly moving particle-observer of the system sees himself for all time as the centre of the system, and sees the system arranged non-homogeneously but with spherical symmetry round himself. He can of course readily assume apparently excentric positions. He simply has to assume a velocity **V** different from the value associated (in any frame) with the given particle near himself. He at once sees himself at a vector-distance **P**—**V**t from the apparent centre of the system, and immediately undergoes an acceleration in the direction of this apparent centre.

Strange though the properties of the system may appear, they involve no mystifying paradoxes. They are all rationally explicable, and the calculations relating them are throughout of the simplest character. They are all capable of being described in terms of the measures of customary physics, and the system can readily be illustrated by a diagram. It is completely free from self-contradiction.

Though the system always occupies, in the experience of any member of it, a finite volume, this volume of particles has no 'velocity through space'. The system possesses a centre of position in the experience of each particle occurring, coinciding with itself, and is 'at rest' as a whole relative to this particle. It is equally at rest relative to any other particle. A particle-observer O considers the system as a whole to have velocity zero in his own reckoning, but he is not thereby entitled to consider it as having a velocity —**V** in the experience of a particle-observer P moving with velocity **V**. In P's experience the

system is again on the whole at rest Every observer has his own space-frame, but no meaning can be assigned to asking what is the space-frame common to the different observers

118. This system achieves in a simple way what has previously been accomplished by the use of a 'curved expanding space' But the resulting body of experiences of every given particle-observer is essentially different from the experiences of an observer in the 'expanding-space' systems, as we shall see In the next chapter we consider the application of this system to the problem of giving a rational account of the universe

ADDITIONAL NOTE

P 97 If universal constants exist, the function G controlling the behaviour of *additional* test particles may be of the form $G(X, \xi)$ instead of simply $G(\xi)$ But the motion of the *given* particles in the simple kinematic system is unaffected

VI

THE EXPANDING UNIVERSE OF NEBULAE

119. THE total number of extra-galactic nebulae capable of being photographed down to the threshold of identification with the world's largest telescopes is very large indeed. Hubble has pointed out (1934) that the nebulae increase with limiting magnitude at such a rate that near the galactic poles they equal the stars in number at about magnitude 21.25, which is about the limit of identification that can be reached with long exposures under good conditions with the 100-inch Mt. Wilson reflector. He estimates the total number at the limit as 75,000,000 nebulae. This may be compared with the estimate of Shapley of a number of nebulae exceeding 300,000 down to magnitude 18.2, which is in approximate agreement with a similar estimate published by Seares in 1925. Of 76,000 nebulae catalogued by Shapley in 1933, he estimated that the greater number lie between 30 million and 100 million light-years distance. He has also estimated them as distributed on the average at the rate of one galaxy or nebula for each cube of side 1.3×10^6 light-years. The corresponding density of matter in space if the material forming the galaxies were smoothed out to a continuous distribution, assuming a mass of 10^9 suns per galaxy (1 sun $= 2\times 10^{33}$ grammes), is 10^{-30} grammes per cm.^{3}, and a similar estimate has been made by Hubble. The true density of matter in space may exceed this estimate several thousand times as the galaxies may possess much larger stellar populations than 10^9 suns per galaxy, and, besides that, allowance must be made for matter in inter-galactic space. The number of stars in our own galaxy has been estimated as of the order of 10^{11} suns, and there is increasing evidence that the average galaxy is comparable with our own. This change alone would increase the mean density of matter in space to the order of 10^{-28} gramme cm.$^{-3}$

The galaxies, though showing appreciable local clustering, are distributed approximately homogeneously in 'our own neighbourhood', say up to 100 million light-years' distance. The actual testing of the homogeneity is a matter of some difficulty, since it depends on the calculation of distances from apparent brightnesses. When certain 'corrections for red-shifts' have been applied, Hubble finds a tendency for the apparent density to increase outwards. But no allowance has

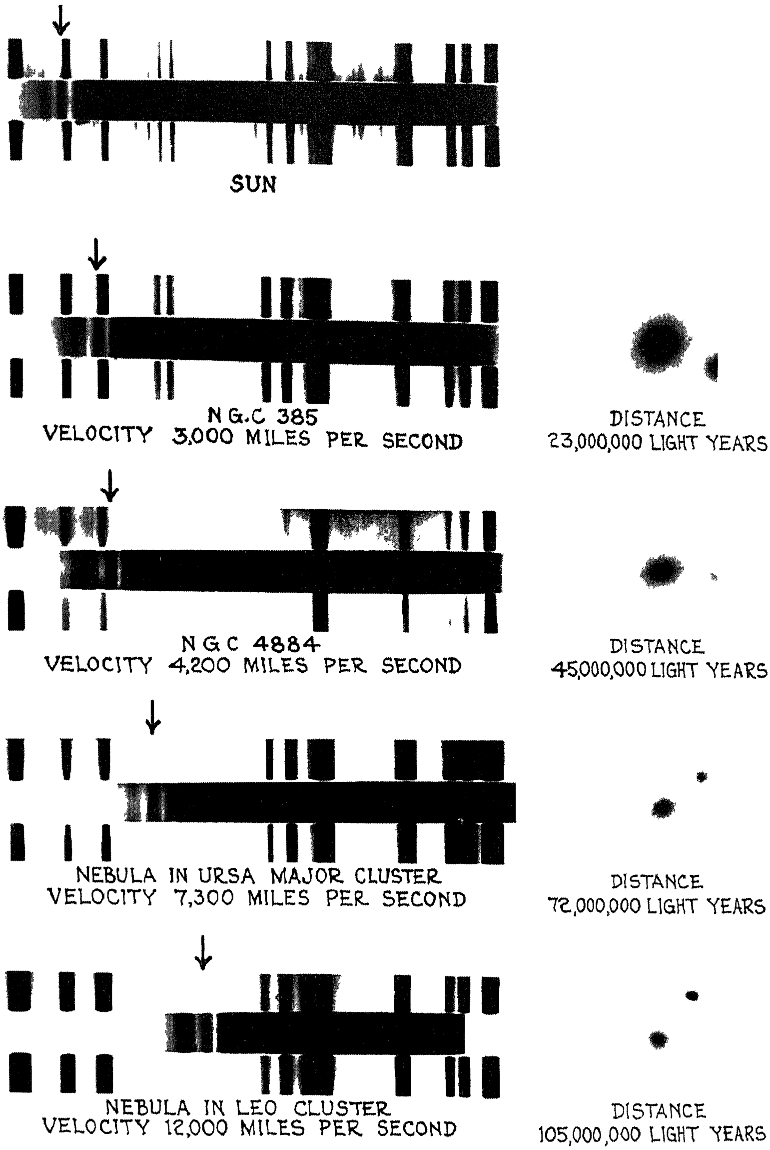
as yet been made for the fact that a photograph of a region of the sky shows the nebulae not at a common instant of time in the experience of the observer making the photograph but at a common epoch of observation The brightnesses of the more distant ones therefore correspond to distances which are smaller than the 'simultaneous' distances by lengths which depend on the speed of recession and on the time of travel of the light to ourselves This effect has been computed in Chapter IV, and gives an apparent density in the photograph exceeding by a factor $1+4v/c$ that for a distribution homogeneous in the local time scale of the observer Hubble's result that the apparent density increases outwards is probably this effect, though it is impossible at present to say whether it is confirmed numerically

All these nebulae are animated with a velocity of recession from ourselves, if we assume that the observed red-shifts in the spectra may be legitimately interpreted as velocities (We have seen that this is the only possible interpretation if the nebulae are equivalent to one another, an effect of difference of gravitational potential would mean that if a nebula A observed a nebula B as possessing a red-shift, A to B would possess a violet shift, and the nebulae would not be equivalent It would be a highly improbable state of affairs that we ourselves should be in a region of the universe of lower gravitational potential than the region of any other nebula) Hubble's law states that the velocities are proportional to the distances, whence it follows that the relative velocity of any two nebulae is proportional to their relative distance, and that the nebulae are all receding not only from ourselves but from one another

120. It is not our object to describe here in detail the fascinating researches now in progress on such questions as the clustering of galaxies, the sizes, populations, and masses of galaxies, the structure of the various objects composing them—stars, Cepheid variables, novae, globular clusters—their internal motions or the existence of clouds of absorbing matter in their vicinities Some of these we shall require to discuss later. For the present we propose to examine some of the fundamental questions raised by the above bald summary

Is the Universe finite or infinite?

121. The first question which presents itself is whether the total number of galaxies that could possibly be observed is finite or infinite Imagine a telescope of very great light-gathering power, and



Spectra of nebulae, showing the shift of spectral lines to the red The arrow indicates the *H* and *K* lines of ionized calcium

Note In this print the spectra have been widened and slightly retouched, mainly to reduce the effect of false lines due to the grain of the negatives The *H* and *K* lines have not been retouched

(*By the courtesy of the Director of the Mount Wilson Observatory*)

let the nebulae visible in it be counted Let the aperture be increased and the count repeated, and so on indefinitely. Then either the total number of nebulae counted tends to a limit or it does not

122. The question of the finiteness or otherwise of the universe has usually been treated as a single question In reality it is two One question is: is the total number of galaxies in the universe finite or infinite? A second question is: is the region of space accessible to observation at any one epoch in the experience of a given observer finite or infinite? The former question is one of fact The second is primarily one of the geometry and coordinates adopted by the observer Assuming he has adopted the flat-space and coordinate conventions as to distance and epoch defined in the present work, which are those of ordinary physics, the second question is also one of fact Are the objects observed found within a finite volume of space at any one epoch or do they extend indefinitely far?

123. Eddington, in an imaginary conversation described in his book *The Expanding Universe,* begins by taking for granted the answer to the first question He says, 'I want you to imagine a system of say a billion stars spread approximately uniformly, so that each star has neighbours on all sides.' A billion, though large, is a finite number Let us investigate the consequences of Eddington's hypothesis

If a billion stars can be observed, they can in principle be measured as to their distances by parallax observations, or their distances can be estimated from their apparent brightnesses Since the number of stars is finite, there will be a greatest distance Select the star at this greatest distance. Then in the field of view of a telescope large enough to observe this star, it will be discretely separate from its neighbours, and will appear in a definite direction Let us move towards the star in this direction, with our telescope Then the distance of the selected star and its neighbours in the field of view must diminish Either the selected star remains the most distant one in the direction of the field of view or other stars must come into view beyond it

124. In the former case we ultimately reach the star in question, and we are then at the confines of the universe of stars in this direction The universe would then possess an accessible boundary or edge On one side of the boundary would be material objects, on the other side

a void, receiving light and other forms of radiation, but containing nothing, radiating nothing back again. This is simply a consequence of the assumed finiteness of number of material objects contained in the universe. From the same hypotheses made, in every direction there would be a finite accessible boundary and the universe would be contained in a closed finite volume, surrounded by emptiness. Being finite in number the objects in the universe would possess a definite centre of position and a definite velocity-centroid, defining an absolute standard of rest and absolute position in space. In infinite space we should have found absolute standards.

This state of affairs contradicts the requirement that every star shall have neighbours on each side, but it may be objected that there is no need to impose this requirement. I think that this objection can be maintained, and that a finite universe surrounded by an infinite void is a conceivable universe. But it is not an intelligible universe, for by the nature of things the question why one part of this void should be populated by matter and the rest be empty is for ever incapable of answer. Astronomers, cosmologists, and other thinkers have invariably rejected this as a possible universe. Einstein has written, 'The stellar universe ought (then) to be a finite island in the infinite ocean of space. This conception is in itself not very satisfactory. It is still less satisfactory because ... such a finite material universe would be destined to become gradually but systematically impoverished.' It is not essential for us to assent to the latter argument, but as our object is to try to understand the universe we should at this point have definitely to admit failure in our quest if we were driven to the island universe theory.

125. Returning to the case of a finite number of stars in the universe, the second alternative was that as we approached the most distant star in the field of view other stars should come into view beyond it. But then retracing our path with an enlarged telescope we could keep these stars in view until we returned to our starting-point. In that case the selected star would *not* be the most distant star visible, and we should have reached a contradiction. The only way out from this contradiction is to suppose that initially, although the total number of stars *in existence* was finite, the total number of stars *visible* was infinite. In that case some stars must be seen twice over, or several times over, and a star of this class would be in view in more than one

direction Let P and Q be two different apparent positions of the same star, at finite distances We could journey towards P, ultimately reaching it, and meanwhile keeping Q under observation in our telescope Q would then be a portion of P, seen from P, and at P we should be viewing ourselves also at Q The same object would then be in two places at once This is a common phenomenon in optics —we call the one entity the object and the other its image This is the 'ghost' theory—that some of the visible stars are mere optical images of real stars But in that case actual observations could be carried out, in principle, which would distinguish images from objects —we could, in principle, journey to each observable entity and ascertain which of them were 'real', which optical illusions We could then discard consideration of the illusions, and we should be left with a finite number of real objects One of these would have a greatest distance, and we could apply all our arguments over again We should be driven back on the island universe theory Alternatively, if all the visible objects were 'real', and the same object could be in two distinct places at once, we should have to give up again all hope of a rational understanding of nature or indeed any rational physics

126 These are the consequences of the hypothesis of a universe containing a finite number of objects There are many other consequences equally unacceptable when we come to consider the velocities of the objects For, as above, the velocities of any finite number of observable objects must be less than the velocity of light They therefore determine a mean velocity, which may be taken to define 'rest', and the meaning of this selection of a preferential frame by the material objects in a featureless space is for ever unintelligible

127 Rejecting, in common with all other investigators, the 'island universe' theory, we are driven to the conclusion that the universe cannot contain a finite number of objects Hence it must contain an infinite number

128. We have reached this conclusion without at any point finding it necessary to discuss whether 'space' is, or is not 'curved', whatever significance such a question may or may not have How does it come about that many investigators, adopting the concept of a closed, finite, curved space for the description of the universe, have reached

the contrary conclusion, namely that the universe contains only a finite number of objects—stars, nebulae, atoms or what not?

The answer to this is twofold First, a mathematical technique undoubtedly exists which describes a finite collection of particles or material objects, each surrounded equally by neighbours, which avoids, or avoids in part, the contradictions we have encountered above But it does this at the cost of *creating* further matter, in time, as fast as it is required This was not the fault or object of the originators of the technique They merely demanded a finite number of objects each equally surrounded by other objects of the system The mathematics they used proved itself equal to the demand, but only by surreptitious creation of fresh objects as fast as they were required to be observed, and observable as distributed round objects already created This avoids the difficulties encountered above, where no creation of objects in the continuing experience of the observer was permitted This result is not generally recognized, but we shall establish it as a rigorous consequence of the current relativistic treatments of cosmology It can and will be shown, in fact, that in what is called the Einstein-de Sitter universe, a telescope, exposed at a given epoch, however large its light-gathering power, would never disclose more than a definite finite number of objects which can be calculated Indefinite increase of aperture, at any given epoch of observation by a given observer, leaves this number unaffected But as the observer gazed through the telescope, the theory in question predicts that he would continually see fresh points of light appear, not previously visible They would appear at first of very small brightness, animated with a velocity of recession close to that of light, then steadily decelerate and for some time become brighter, ultimately fading away again as the distance increased Similar phenomena are predicted on other current relativistic models Each luminous object—such is the prediction of these theories—appears in the field of view at a definite epoch in the experience of the observer concerned, having been previously invisible Beyond any given object other objects appear in turn in unending sequence, so that each object as soon as 'born' is immediately surrounded by further objects The total number of objects tends to infinity as the epoch of observation increases, although finite at any given epoch of observation This state of affairs is indistinguishable from a continual creation, in time, of distant objects

I shall demonstrate this in detail in Chapter XVII The mathematics, set in effect an impossible task, plays a trick on the investigator It acts like a Delphic oracle, and unwilling to acknowledge defeat has recourse to the device of creation to satisfy the demands placed on it

The other half of the answer is that the arguments which have been used to bolster up a belief in the possibility of a finite universe are fallacious False analogies have been employed Eddington, for example, in the imaginary dialogue already mentioned, makes the defender of curved space remark 'I could arrange a billion people on the surface of the earth so that each of them has neighbours on all sides, and no question of a boundary arises I only want you to do the same with the stars' But the people have a boundary—the curved surface of the earth We cannot arrange a billion people on the earth to have neighbours on all sides, in the arrangement contemplated, they have no neighbours above them or below them The stellar aggregate contains no entity corresponding to the material surface of the earth The phenomena which compel our belief in a curved earth—such as the bit-by-bit disappearance of a receding ship—have no counterpart amongst the stars If indeed we supposed that objects could only be seen by rays which hugged the curved surface of the earth (e g radio waves) there would then be no disappearance of a receding object, but instead we should see each object, in principle, in two antipodal places at the same time, and we should come back to the state of affairs of § 125 Thus we secure a finite visible population in an unbounded space only by introducing twinning In the universe favoured by Eddington it is impossible to distinguish between the antipodal objects

I am not criticizing the use of a conceptual curved space for certain purposes Riemannian geometry is one of the most splendid tools at the disposal of the mathematician The reasoning of those mathematicians who have used curved space as a means of description of possible universes is flawless I am only concerned to argue that a universe of a finite number of particles involves difficulties, and that such difficulties are necessarily encountered by any theory which adopts a closed finite space filled with matter at a finite density The reasoning of those who without mathematics have attempted to persuade readers of the 'reasonableness' of a curvature of space, homogeneously filled, as the only means of description of the world is

not flawless Once we have recognized that geometry is arbitrary—the position adopted in Chapter I—a curved space is equally admissible, of course, with a flat space But we must be careful to ascertain what observations it implies, and how these would be rearranged by an observer using flat space

Eddington has written: 'Few scientific men nowadays would reject spherical space as impossible ' This is a deplorably loose statement, but he seems to mean that there may be some ultimate character of an entity he never defines, namely 'physical space', which may be denoted by the term 'spherical' He asks what is the supposed disadvantage of spherical space I have pointed out above the contradictions to which we are led by assuming a universe of a finite number of particles, which is the consequence of adopting spherical space He goes on to argue 'A closed system of galaxies requires a closed space If such a system expands, it requires an expanding space This can be seen at once from the analogy that we have already used, viz , human beings distributed evenly over the surface of the earth, clearly they cannot scatter from one another unless the earth's surface expands ' Of these statements the first is question-begging It assumes that the system of the galaxies is closed The second statement is argued by an analogy we have shown to be false The last statement involves a further false analogy, for it introduces a material entity, the curved surface of the earth, as a subject for the verb 'to expand', whilst the phrase 'expanding space' only has a meaning if we invent a something possessing a density or linear dimensions which can change in time, for example an ether If the dimensions that are expanding are the distances between galaxies, for which Eddington has already admitted a meaning, let us say so directly Let us not invent some undefined entity, physical space, merely to be able to assert the otherwise meaningless proposition 'space expands'

The matter could be argued at much greater length In this book I am more concerned with constructive results than with tilting at obsolescent modes of thought I am content to conclude that a universe containing a finite number of objects would be irrational, and to examine now the contrary thesis

Properties of a universe containing an infinity of objects

129. We first notice that the proposition that the universe described in astronomy contains an infinite number of nebulae can never be

established by observation. All observation can do is to count the number of objects seen, and this number must necessarily be finite. It is important to realize that one of the most fundamental questions that can be asked about the totality of things cannot be answered by pure observation, but must be answered if at all by an admixture of reasoning. We cannot ever answer by observation the question whether the universe contains an infinite number of observable objects.

Observation has in fact disclosed an increasing total number of nebulae with every increase of limiting magnitude. If the total number of nebulae observable is infinite, observation should disclose nebular objects as close together as can be resolved or separated, together with a background of irresolvable objects of finite total luminosity. There are no observations which contradict this. No limit has been found to the closeness together of nebulous objects in photographs as far as identification is possible, and there is some evidence of a luminous background. A universe containing an infinite number of nebulae is compatible with modern observation.

130 An infinite population of nebulae immediately removes velocity difficulties. There will now be no fastest nebula, instead, nebulae can exist with all speeds extending up to, but never reaching, that of light. The swifter they move, the more the recession will reduce the apparent brightness, thus making possible a finite background luminosity. Further, with an infinity of nebulae, there is not necessarily any unique velocity-centroid, and no absolute standard of 'rest' is imposed.

131 This leaves untouched the question of whether the volume occupied, by the infinity of nebulae, is finite or infinite. The maximum observable distance may be finite or infinite. If it is infinite, the population of nebulae could be spread discretely through infinite space. If it is finite, the nebulae must form an 'open' set of points, crowded together towards but wholly within an inaccessible boundary. (For if the maximum distance were actually *attained* by a nebula moving with some speed less than that of light, and the boundary were thus in principle accessible, the difficulties of the island universe theory would again be encountered.) Now Hubble's law of proportionality of velocity to distance is not only observed to be fulfilled as far as observations yet go but is given immediately by all current relativistic theories, by the most elementary non-relativistic kinematic considerations, and by the relativistic kinematic model

Assuming, as we may, its probably exact satisfaction when flat space is adopted for any given observer, for distances and velocities reckoned by that observer, it predicts that the velocity of light would be attained in a finite distance, and therefore that all observable receding nebulae must be inside a certain limiting distance at any one given epoch The volume occupied by the universe is therefore probably finite although the total population is infinite In that case the nebulae, infinite in number, are representable as an open set of points, crowded together at any given epoch towards the boundary, which would, however, be inaccessible and moving with the speed of light The nebulae themselves, near the boundary, would be subject to an extreme Lorentz contraction, in the view of any observer, and would indeed in the limit appear squashed flat in the direction of their motion There is therefore abundance of 'room' for them, infinitely dense though they be in distribution near the inaccessible expanding boundary

We have reached these conclusions simply by rejecting the hypothesis of the finiteness of number of nebulae in the universe, on account of its difficulties, and by adding to the alternative hypothesis of an infinity in number of nebulae an extrapolation of Hubble's law of proportionality of velocity to distance.

Comparison with the simple kinematic system

132 Now compare these conclusions with the properties of the system constructed in Chapter V, whose construction was independent of the arguments of the present chapter. This system was not constructed so as to satisfy Hubble's law, or to contain all velocities up to that of light, or to contain an infinite number of particles. It possessed all these properties, but they arose as consequences of the consideration of a system of particles in uniform velocity satisfying the cosmological principle Hubble's law arose as the limiting form of the position-velocity correlation naturally assumed, in the limit, as a consequence of the natural motion of a cluster of particles in uniform relative motion The existence of all speeds up to that of light, the form of their distribution, and the infinity in total number of particles arose as the unique way in which the system of particles, constrained to move uniformly, could satisfy the cosmological principle In that case every particle was equally and (locally) uniformly surrounded by other particles, the boundary was inaccessible, and the particles were

crowded together towards the boundary as locus of limiting points Lastly we saw that the constraints could be removed, and that every given particle would then in fact move uniformly as a free particle, but that an additional freely projected test-particle would be subject to an acceleration in a definite direction which could be calculated and interpreted physically

133 The kinematic system constructed theoretically in Chapter V is therefore fully suitable, in respect of the properties mentioned, as a representation of the universe of nebulae I definitely propose this system as a model of the universe as disclosed in astronomy, a model in which the nebulae are represented as particles

The 'sample' principle

134. In his Halley Lecture, 1934, Hubble has written 'Observations give not the slightest hint of a super-system of nebulae Hence, for purposes of speculation, we may invoke the principle of the Uniformity of Nature, and suppose that any other equal portion of the universe, chosen at random, will exhibit the same general characteristics As a working hypothesis, serviceable until it leads to contradictions, we may venture the assumption that the realm of the nebulae *is* the universe—that the Observable Region is a fair sample, and that the nature of the universe may be inferred from the observed characteristics of the sample '

This assumption, the 'sample principle', was of course fully in my mind in writing Chapter III, but I carefully refrained from using it, because it is not by itself powerful enough to afford a basis for the construction of a cosmology. A system which satisfies the cosmological principle necessarily satisfies the sample principle For if the description of the whole universe is the same, in space-behaviour and time-behaviour, from all points of view defined by all particle-observers occurring in the system, then description of the *local* samples accessible to *local* observation will be identical The cosmological principle requires something more than this—it requires that not merely the *local* samples be described identically, but that also the *whole* system be described identically from different points of view The world-wide description must be independent of origin

Relation of the sample principle to the cosmological principle

135. The relation of the sample principle to the cosmological principle may be seen as follows Suppose we survey and describe, from a

particle O, a local sample surrounding O Let P be a particle different from O Let the system be now surveyed and described from P. From P a new local sample will be visible, which will partly overlap the old sample but will contain new particles P', not practically observable from O The principle mentioned by Hubble demands that the local survey from P coincide with the local survey from O This allows us to determine the characteristics of particles P' not observationally accessible from O The sample principle is a rule of extrapolation, permitting extrapolation to particles P', as seen from P, of properties observed for particles P as seen from O Let the process be repeated from P', and so on indefinitely Clearly the process never comes to an end The cosmological principle is now the statement that if at any stage of the process, the whole system so far constructed be surveyed from O and from any other particle P, then in so far as the fields surveyed overlap, the descriptions of the characteristics of portions constructed are indistinguishable Eventually the cosmological principle constructs a system such that if the *totality* be surveyed from O, it will be described in the same way as the totality surveyed from P The sample principle is just the cosmological principle applied locally, and the cosmological principle is the limit of the result of applying the sample principle an infinite number of times

136. The cosmological principle does not prescribe *a priori* whether the total population is infinite or not, whether it is locally homogeneous or not, whether it possesses world-wide homogeneity or not, whether it is included in a finite volume or not It leaves the answers to these questions to be inferred by application of the principle When we begin our analysis on the basis of the cosmological principle, we do not know what the answers will be We eventually find as consequences of our application of the principle, that the whole system will be infinite in population, will be locally homogeneous, will possess a density-distribution increasing outwards in the experience of any observer, and will occupy a finite volume in the experience of any observer

The cosmological principle as an expectation

137 In Chapter III, I considered the cosmological principle as affording a method of constructing systems which would afford standards of comparison It was the simplest substitute for the unworkable concept of homogeneity The sample principle, on the other hand, is

the statement of an *expectation* If the sample principle is everywhere obeyed, both now and for any future improvements in the means of observation which enlarge the boundaries of the locally accessible region, then the cosmological principle must also be obeyed For the cosmological principle, considered as an expectation, is simply the statement that no possible increase of telescopic power can lead to a violation of the sample principle The region accessible to observation is expected to be a fair sample however large it may be, indeed until the 'sample' embraces the whole system This 'expectation' is of a metaphysical character, as already explained It is an extension of the principle of the uniformity of nature from the independence of *laws of nature* as to locality and epoch to the independence of the *structure* of the universe as to point of observation, where the word structure includes the complete history of the motions If the sample principle is adopted as a working hypothesis, it becomes free from contradiction only when broadened into the cosmological principle The expectation that the universe of nature, idealized to a collection of particles in motion continuously distributed, must satisfy the cosmological principle, I originally called 'the extended principle of relativity', but I do not now consider the name well chosen The two aspects of the cosmological principle—its aspect as a definition, or rule of selection, as developed in Chapter III, and its aspect as an expectation, or rule of extrapolation, as here developed—must be carefully distinguished

Why extrapolate?

138 There will be some readers who ask at this stage 'Why extrapolate at all? Why not be content with the domain accessible to observation, and confine ourselves to cataloguing its characteristics?' To such readers I reply that they have no business to be reading this book If they are only interested in the domain accessible to observation, they need not go beyond treatises on astronomy But to those who are interested in vaster questions, the cosmological problem immediately suggests itself What if we *do* extrapolate? And extrapolate again and again? Do we construct in the limit a system of an infinite or a finite number of particles? Does the constructed system surveyed not only partially, as an extrapolated chunk might be surveyed, but surveyed as a whole, occupy an infinite volume or a finite volume? The cosmological problem is the problem of the

structure and history of the totality of things. And the answer to the cosmological problem is best sought by constructing systems satisfying the cosmological principle. The cosmological principle is the only possible rule of extrapolation which avoids attaching preferential characteristics to the portion of the universe already surveyed. It is the grandest rule of extrapolation that can be imagined.

Extrapolation of observations in relation to the cosmological principle

139. We now note a surprising fact. The cosmological principle *might* have been applied empirically, by cataloguing the observed characteristics of the portion of the universe already surveyed, and then extrapolating them. This would be a perfectly fair method of procedure, though difficult of execution. Actually we have proceeded quite differently. We chose to consider two abstract problems, the kinematics of a cluster of particles constrained to be in uniform relative motion, and the local kinematics of any system obeying the cosmological principle locally (see Chap. IV). By combination, generalization, and idealization of them we constructed, without recourse to observation, a system rigorously satisfying the cosmological principle, throughout its domain and throughout its history, without constraints. We now find that this reproduces, within the limits of observation, many of the characteristics of the observable system of nebulae. It follows that we have got the same results as if we had actually extrapolated the observations.

This is not to say that there is only one way of extrapolating the observations. The accelerations or decelerations of the nebulae, if they exist, have not been observed. I do not know yet whether systems of accelerated fundamental particles satisfying the cosmological principle can be constructed kinematically. But in the absence of knowledge of the nebular accelerations, the least committal assumption is to work on the hypothesis that they are zero. We should then be led inevitably to the kinematic system of Chapter V. This may be considered accordingly the simplest possible system that can be obtained by repeated application of the sample principle, or by means of the cosmological principle.

Predicted properties of the nebulae

140. It is a good plan in research always to consider particular cases first. It is easier to see what is happening. It is therefore well worth while to examine the kinematic solution, even if later research proved

it to be only a particular case, to see what it suggests about the universe as a whole According to it the following properties should hold good for the system of the nebulae

(1) The nebulae at distance r at epoch t in the experience of an observer situated on one of them possess velocity r/t directed radially outwards

(2) (*World-map*) The number of nebulae between distance r and $r+dr$ at epoch t is proportional to

$$\frac{tr^2\,dr}{(t^2-r^2/c^2)^2}$$

(3) (*World-picture*) The number of nebulae viewed *at epoch of observation* t between apparent distances r_1 and r_1+dr_1 is proportional to

$$\frac{r_1^2\,dr_1}{t(t-2r_1/c)^2}$$

(4) The density is locally constant in space, in the world-map, and decreases in time as the inverse cube of the epoch It increases outwards, but the effect for small distances is of the second order

(5) The density in the world-picture decreases locally in time as the inverse cube of the time, but increases outwards locally according to the proportionality $n \propto 1+4\frac{r_1}{ct} = 1+4\frac{v_1}{c}$, where v_1 is the observed velocity

(6) (*World-picture*) The relation between observed velocity v and observed distance r_1 *in a photograph taken at time t is*

$$v_1 = \frac{r_1}{t-r_1/c} \sim \frac{r_1}{t}\left(1+\frac{r_1}{ct}\right)$$

The *observed* velocity should therefore show an increase with *observed* distance, faster than the first power of the observed distance

Properties (5) and (6) should serve ultimately to establish the truth of the interpretation of the red-shifts as recession-velocities

(7) The total number of nebulae is infinite

(8) The recession velocities range up to c

(9) The distance of any nebula is less than ct

(10) The density tends to become indefinitely great at the distance ct

(11) The nebulae at distances just less than ct are almost invisible

(12) There is a finite non-zero background intensity, below any given limit of identification V_1, proportional to $1-V_1/c$ (See Note 7)

(13) The nebulae visible at very great distances, at the events we now perceive, possess a very early local time-reckoning t' as compared with our epoch of observation t, and so should be in a relatively earlier stage of evolution

(14) Nebulae at great distances should show a strong Lorentz contraction in the line of sight

Further remarks on space and geometry

141. The system we are discussing has been described in terms of the flat space and ordinary time of any observer situated on any arbitrary nebula of the system, which may be itself considered as the centre of the system This solves a problem previously only solved by the use of curved or flat expanding spaces Eddington has remarked 'It is well known that the assumption of flat physical space leads to very serious theoretical and logical difficulties ' He is here quite mistaken The system we have described is free from all internal contradictions and free from unacceptable paradoxical consequences It is easily pictured and represented in a diagram Eddington has been led to this error by his view as to the nature of space He has written 'There is of course no reason for supposing space to be flat unless our observations show it to be flat, and there is no reason why we should be able to picture or describe the system in flat space if it is not in flat space'—and he proceeds once again to reason by analogy, talking of square pegs in round holes The proof of the pudding is in the eating—we have shown that the objects for which Eddington required a curved space can be achieved by flat space Eddington's view as to the nature of space is diametrically opposed to that of Poincaré, who (as mentioned in Chapter I) held that no observation can distinguish whether space is Euclidian or non-Euclidian, simply because there is no such thing as physical space, but only the space, of our arbitrary choice, which we create and use for purposes of description

142 Doubtless the finite domain of flat Euclidian space occupied by the infinity of nebulae in this model, namely the sphere of radius ct, could be represented in a closed curved space. But the density would not then be constant It is particularly to be noticed that I have nowhere contended *both* that the geometry assumed may be adopted arbitrarily *and also* the density in this geometry may be

taken to be uniform We cannot simultaneously choose the space and choose the density-distribution We have chosen our space or geometry arbitrarily, but we have left the space-filling particle-distribution to determine itself on the basis of the cosmological principle We make no *a priori* decision that the density shall be uniform, and we have found in fact a density-distribution locally homogeneous but increasing ultimately outwards in the experience of any arbitrary particle-observer of the system This is one of the main advantages of resting the whole investigation on the basis of finding systems satisfying the cosmological principle, instead of prescribing *a priori* a homogeneous distribution together with a geometry

143. It follows immediately from Poincaré's view as to the arbitrariness of geometry that no knowledge whatever concerning the universe can be obtained by making assumptions about its geometry We must discuss *phenomena*, not geometry. Actually when writers 'adopt a metric' for the universe, they are doing two things (1) choosing a geometry, (2) assuming that in this geometry the free paths are given by geodesics, $\delta \int ds = 0$ They thus posit a set of *motions* But this still conveys no information as to phenomena, for the distribution of material in the presence of which these are the free paths is still left entirely undetermined It only becomes determinate when some further assumption is made which allows the distribution of matter present to be inferred from the adopted geometry This assumption is carried out by (3) adopting a set of 'field equations' As these equations are still arbitrary as to an unknown constant, the cosmical constant λ, the distribution of matter present is still largely arbitrary (To say that λ is small is to make an empirical appeal to quantitative observations) We have constructed a method which abolishes these three consecutive assumptions by recognizing at the outset the arbitrariness of geometry as a means of description, and choosing arbitrarily the space commonly used in physics, we have then simply sought distributions of matter-in-motion satisfying, in this adopted space, the cosmological principle, taking care to supply first the rules for passing from coordinates in one observer's space and Newtonian time to those used by an equivalent observer, equivalent observers being supposed present as part of the definition of a system satisfying the cosmological principle It is particularly to be noted that we discuss matter and motion simultaneously. The demand that the

system is to satisfy the cosmological principle immediately imposes the distribution of the particle-observers given their equivalent motions We begin by discussing observable phenomena, not geometry

Independence of a theory of gravitation

144. A further advantage of our model of the universe is that its validity as a possible model does not depend on any pre-stated theory or law of gravitation It is a possible set of nebular motions whether or not there be a specific law of gravitation, the motions must continue to conform to the rules given whatever (relativistic) form of 'the law of gravitation' be adopted Thus the model is independent of whether there is or is not a cosmical constant λ

Determination of local density

145 The system we have described contains one arbitrary constant, the constant multiplier B Thus there is a manifold of possible solutions corresponding to one arbitrary variable, all possible distributions being related by simple superposition The degree of arbitrariness, namely dependence on a single arbitrary constant, is the same as for the relativistic solutions, though they further involve the 'cosmical constant' λ We have established the relation

$$\tfrac{4}{3}\pi Bm = -\frac{c^3 t}{\gamma} G(1),$$

where m is the mass of a 'particle' or nebula, but we have not been able to predict the local density from the expansion Nor can the general relativity theories without special assumption If, however, we assume that $G(1) \sim -1$, the local mass-density is given by

$$\rho_0 = \frac{Bm}{c^3 t^3} \sim \frac{3}{4\pi\gamma t^2} = 10^{-27} \text{ gramme cm}^{-3}$$

This is not incompatible with the present estimates of Shapley and Hubble of a density of 10^{-30} gramme cm $^{-3}$, allowing for invisible matter in space and for the probable as-yet-unobserved extensions of the galaxies This shows that the absolute value of $G(1)$ is not far removed from unity But the important thing is that essentially the density predicted by the theory as so far developed is arbitrary, depending on the arbitrary constant B

What is 'outside' the system?

146 The reader will naturally ask, if the system occupies, in the observer's space, an expanding but always finite sphere of radius ct, what is there outside this sphere? Do not the receding nebulae require external space in which to expand?

The answer is that this question is meaningless The background of unobservable nebulae—observable in principle down to any assigned threshold brightness but always including infinitely many nebulae arbitrarily faint below this threshold—this background is in technical language 'everywhere dense' It contains no holes In every arbitrary direction there is some nebula, and the set of such combine to form a continuous background There is no direction in which there is any 'window into outer space' We may if we like *imagine* an outer space But we can never by any conceivable method observe any object in this space, for there is in every direction a curtain of faint nebulae obscuring the vision If we did make the hypothesis that this 'outer space' existed, then if it contained no objects it would never make the slightest difference to any observation whether we made the hypothesis or not Consequently the hypothesis is not a significant one It is immaterial whether the space exists *a priori* or whether the space is created by the system as it expands—immaterial because either view has no verifiable consequences, and our whole philosophy is that an unverifiable proposition† is meaningless If, on the other hand, we made the hypothesis that such space existed and contained material objects which were ultimately overtaken by the expanding frontier and appeared inside the nebulae, then the object would suddenly appear to observation as a new creation, it would in fact *be* a new creation, for our outlook implies that the assertion of the 'creation' of an object is indistinguishable from the assertion that it did not previously exist to observation To say that an object exists but cannot be observed is meaningless the proposition that such an object 'exists' has no content, because it is not verifiable In fact, an object could not be overtaken by the expanding frontier without colliding with some nebula, a member of the expanding system, and *ex hypothesi* a nebula beyond any nebula accessible to observation, for there are always an infinity of nebulae, everywhere dense in the sky, beyond any accessible nebula Consequently the nebula undergoing the collision would be necessarily unobservable, and the event

† That is, concerning the world of nature

of the collision would be unobservable. In other words it is logically self-contradictory to posit any 'swimming into our ken' of objects previously beyond the expanding frontier, the structure of the system itself forbids any such occurrence. Mathematically, we can and shall describe the 'contents of space' outside the frontier of the expanding system. But as would be expected, such objects never come, and never can come, into interaction with observable members of the given system. They may, mathematically described, *tend* to be swept up by the expanding frontier, but they never penetrate it. Penetration would be a logical self-inconsistency of description of the whole system.

147. These considerations solve at one blow the question of whether space is or is not infinite. Infinite space is in daily use by the mathematician. The simplest problems of geometry demand the consideration of infinite space. But such a space is the space of pure mathematics, not the space of physics. Physics, which demands verifiable answers to its questions, must ask: can we observe material objects at arbitrarily large distances? This is the only meaning that can be given to the question: is space infinite? Our answer is: no! The measured distance of any material object, at any epoch of observation, is not only finite but less than an assignable upper bound, ct. And whether 'space exists' beyond observable objects is not a question with a content. The reader is at perfect liberty to imagine such space, such extending emptiness, and to people it if he wishes. He can conceive of independent creations, spheres of activity, located in these inaccessible and unobservable domains—put there whatever he likes. But his speculations are for ever unverifiable, for the existence of such creations or entities can never affect any conceivable observation he can make.

148. It is quite illegitimate for the observer to imagine himself viewing the system from outside. Our attainable knowledge of the universe is the limiting sum of all possible observations made from inside it. The hypothesis that we can make propositions about the appearance of the system to a 'Cosmic Being' (Eddington's phrase) is not merely ridiculous or impious, it is self-contradictory.

Differences from 'general relativity' models

149. The kinematic system here put forward differs from those developed in current relativistic cosmology in the four respects: (1) it

exhibits at every epoch of observation an indefinitely great number of objects accessible to observation, (2) it gives a continuous luminous background of non-zero intensity, (3) it involves no creation of material in the experience of the observer, (4) it assigns zero acceleration to the idealized particles, and thus to the actual nebulae no systematic accelerations on the average

With regard to (1) and (3), all other relativistic models appear to describe a universe containing a finite number of particles accessible to observation at any one epoch but increasing in number as the epoch of observation advances This will be demonstrated later by actual calculation of the number of nebulae capable of being observed by a given observer at any epoch of the observer's experience According to certain of the other relativistic models, particles appear in the field of view, at first with zero luminosity, the luminosity then increasing and ultimately diminishing Each object at the moment of first just-visibility has a recession velocity equal to the velocity of light, and then decelerates Its deceleration at first increases its visibility faster than its recession decreases it, but the object ultimately fades away as the distance effect gains the upper hand, the velocity or red-shift approaching a constant Meanwhile the number of objects visible increases to infinity, as the epoch advances The hyperbolic-space systems of general relativity tend in the limit, in the observer's experience, to coincidence with the kinematic system, as will be shown But a hyperbolic system can at any epoch be sharply distinguished, in principle, from any kinematic system For a hyperbolic system can always be in principle resolved, in the observer's view, into a finite number of discrete luminous objects, and there is no continuous background, though creation of fresh luminous objects continually occurs, whilst any kinematic system possesses a continuous luminous background

The intensity of the luminous background for a kinematic system is evaluated and discussed in Note 7

VII

CREATION

150 To suppose that a material object can be created in our own experience is contrary to all observation. To say that a particle or luminous object appears suddenly in the field of view, having been previously unobservable whatever the telescopic power, is tantamount to saying that it has been created in our experience. For if it existed before, why was it not observed or observable? Thus the systems of current general relativity imply that we can *watch* the event of creation. We can say if we like that the light, emitted since creation, has only just had time to reach the observer. But this is none the less 'creation in our experience'. In the kinematic model, on the other hand, there is no 'creation in experience'. At whatever epoch an object is viewed, the event has a positive finite non-zero local epoch. And every object, being always in the field of view, has always been in the state of having been created throughout the realizable duration of observation.

In my view, we must accordingly reject the systems of current relativistic cosmology as violating the fundamental fact of experience that material is not created in our experience. It is true that the kinematic system can only be described for $t > 0$, its contents are only observable for $t > 0$. As t is traced back to smaller and smaller values, the system shrinks in dimensions, in the experience of the observer concerned, and *in the limit* $t \to 0$ it approximates to a point. We may say if we like that the complete contents of the system were created once and for all at $t = 0$. The system must *have been* created in order to be observed at all. But $t = 0$ is not an epoch any observer can experience. It is 'pre-experiential', to coin a word. *That* creation is an 'ultimate irrationality', to use Whitehead's phrase. There, if we like, we may trace the finger of God in the divine act of creation. But it is in no *man's* experience. This creation before experience is a totally different thing from the creation occurring in the general relativity models, where particles are created, *in time,* as fast as they are required, outside the last visible particle, to ensure the centrality of each particle in the field. Of course, in our kinematic model we do not *posit* an act of pre-experiential creation. The system is taken as satisfying the cosmological principle, and as capable of being

observed, and we then *infer* the existence of a natural zero of time, which possesses, as we have shown, to an overwhelming degree of probability the properties of an epoch of creation Once and for all, the system was created Unlike the relativistic cosmological models, it has then no further recourse to creation

151 Though creation in the kinematic model is essentially a thing of the past in every observer's experience, his experiences of distant, swiftly moving particles include events arbitrarily near to the act of creation The pre-experiential singularity in density at $t = 0$ in a given observer's history is the precise counterpart of the singularity in density in his present experience at the limiting distance ct Events occurring in our 'now' on distant particles whose distance is just less than ct possess very small values for the local epoch. An observer at these distant particles, communicating his experiences with the light he sends, now being received by us, assigns a very small value to his estimate of the radius of the system at these events, and a very small value to his estimate of the then age of the universe His 'age' estimate is in fact $t' = t(1-V^2/c^2)^{\frac{1}{2}}$, which tends to zero as $V \to c$ Thus by sufficient increase of telescopic power we can in principle see events occurring as little after creation as we please The high density at $r \sim ct$ observed by us, according to the model, corresponds exactly to the high density in our past experience, near ourselves, at $t \sim 0$ But in the limit $V = c$, for which $t' = 0$, the nebulae are invisible, owing to recession with the speed of light, so that whilst we might be inclined to say that even in the kinematic model the event of creation is now being enacted at the distance ct, the proposition is really devoid of content, no observations can verify it

The reader who is not at home in the mathematical theory of limits may feel that we are dealing in fine-spun verbal distinctions Actually there is a fundamental difference between saying that creation occurs in our present at the confines of the universe (the demand made by the general relativity models) and that creation is only just a thing of the past, in our present, for nebulae, in principle observable, as close as we please to the confines We are only able to draw this distinction because our system contains an infinite number of nebulae For, in a system of a finite number of members, in which creation is not occurring, either there is a member moving with the speed of light,

and therefore invisible and unobservable, or there is a fastest member moving with a definite velocity less than that of light. The systems of relativistic cosmology† always contain a last-created member, namely that member, created with the speed of light, which has just decelerated into visibility.

Astronomical consequences

152. This all-but-creation at great distances which is a result of our model has several interesting astronomical consequences. The state of evolution of a nebula can be judged to some extent from its structure —the degree of amorphousness or discreteness in its make-up, the development of its spiral arms, etc. It follows, since local time, in a given photograph, decreases with advancing velocity, that the fainter, more distant, more swiftly moving nebulae should show successively earlier stages of evolution. It is thus possible to hope ultimately for definite observational evidence as to the main lines of nebular evolution. For a distance-sequence should statistically be an evolutionary sequence. In the idealized system, events at ourselves have a more advanced local time than any other event we can observe, for always $t > t'$. Hence our galactic system should be in a more advanced state of evolution than any other nebula we can observe. This is in accordance with the general evidence that our galactic system is a late-type spiral. Every other observer, on every other nebula, should statistically have the same experience. Every particle-observer of the system will see himself as the 'oldest inhabitant' of the system. It follows also, from our earlier calculations, that foreknowledge of the future of the evolutionary trend *from observation* is impossible. We can never hope to observe any more advanced evolutionary stage than our own. But it follows also that the whole past history of ourselves is written in the heavens for us to read, in the various evolutionary stages exposed to our experience. For if the nebulae are, at least statistically, equivalent, then the evolutionary history, recorded in its own local time, t', of any one nebula is the same as that of any other. All the local histories are statistically superposable. Hence the different evolutionary stages at present exposed to view should be the exact counterparts of previous evolutionary stages in our history.

Take any event marking an evolutionary stage. We might take

† In these as in other comments, I am contemplating those general relativity systems for which λ has been taken to be zero.

for example the formation of spiral arms, or the development of a given degree of discreteness in the nebular structure, or the formation of solar systems, or the first appearance of life, or the emergence of man, or the discovery of fire, or the invention of printing. Each will have a definite epoch reckoned from the natural zero of time, that is to say, from creation. Statistically, the same event should occur on each nebular system at the same local time. Thus, if we wish to observe an event whose evolutionary epoch is t', from the formula

$$t' = t(1-V^2/c^2)^{\frac{1}{2}}$$

we compute $V = c(1-t'^2/t^2)^{\frac{1}{2}}$ and examine nebulae moving with the speed V. Statistically, these should show the occurrence of the event in question.

The future of the universe

153. We shall obtain later striking results concerning the probable mean *structure* of nebular systems. Here we are considering each nebula simply as an idealized particle. Even within this limit we can say something about the future of the universe. We can say, in fact, that though each local system evolves, decays, and possibly in some sense grows cold and dies, yet the universe lives for ever, that it knows no death in time. For however large is t, we can always find a nebula so swiftly moving that events on it, observed by us now, are arbitrarily early in time. However small is t' and however large is t, the answer is simply $V = c(1-t'^2/t^2)^{\frac{1}{2}}$. However old we ourselves may be, our experiences always contain events as early as we please in local time. These events we place near the expanding frontier of the universe. For observers at them, experiencing them, the universe to them has hardly embarked on its career of evolution. For them, time is still young. They are experiencing the early history of the universe. The infinite number of nebulae contained in the universe include an unending sequence of arbitrarily young experiences. There, near the confines of the visible universe, is the coming generation. There, we find the sons of the morning. The confines of the universe constitute, as it were, a layer of cosmic protoplasm, expectant of evolution. As I have elsewhere quoted, there the world is

For ever piping songs for ever new,
For ever panting and for ever young.

In our vision, time there stands still, like Grantchester clock in Rupert

Brooke's poem The epoch at which it stands still is not 'ten to three', but it is equally definite, it is the epoch close to creation

Our universe is no passing thing Creation was once, is always The miracle of creation, the unique event of the calling into being of things describable, never repeats itself, but there are always places where it is only just an affair of the past Once started, the system goes on for ever Each constituent has a temporal experience—each observer lives in time but the universe as a whole knows no time history It is the same yesterday, to-day, and for ever. 'Time like an ever-rolling stream bears all its sons away ' Death and decay in our midst, for us, but for the world, immortality The totality of things created knows no terminus in time, no decay, no asymptotic strangulation of the surge of life Always there is a future vista, and since we need suppose no exact parallelism of the evolutionary trends on any two nebulae, there are unending opportunities for variety of local experiences The world is an aggregate of experiences, each experience experienced but once by each inhabitant, but the experiences themselves go on eternally We need appeal to no cycles of rejuvenescence The world ever sows at its frontier the seeds for its own future Each individual nebula reaps the harvest of its own experiences, and passes to the winter of decay But ever anew the seasons recur There at the confines of the visible universe, at the world's inaccessible edge, the music of the spheres is the song of a new dawn, the dawn of the world's perpetual birthday

* * * * * * * * *

Creation and deity

154 The view of the universe I am advocating is no irrational view The system to which we have likened the universe is an intelligible system It contains no irrationalities save the one supreme irrationality of creation—an irrationality indeed to physics, but not necessarily to metaphysics The rest is simply a development of the self-consistency of observations We can give an intelligible description of the universe from every different possible point of view in it, at every epoch capable of being experienced If I am asked for a

description of the universe independent of any observer, I reply that the question is impious Even to conceive of such a question is to imagine as possible to man the assumption of the attributes of deity Yet the reader will already have reflected that the universe I have described is indistinguishable from what we should be disposed to posit for the handiwork of a divine creator as visible to man I have posited no God But I find creation, and immortality (not of man but of the system created), that absence of absolute standards of rest or position which we can call divine extra-territoriality, frontierlessness, infinity of content All these we can describe, with the aid of the technical modes of expression developed by mathematics, from the point of view of the things created To take the one step further, to go beyond description, to suppose an intelligence or activity other than the totality of things experienced, is to add God to the scheme We seem bound to do so For the main part, the physicist gains nothing by the addition of God The system, created, goes on of itself without further acts of creation But the creation was not a creation in absolute time, for no such statement has a meaning It was only a creation in *each observer's* time, and no two observers, in any cross-section of events simultaneous for any given observer, assign the same 'time ago' to the event of creation But creation demands a First Cause And once we have added a First Cause, our system sets no limit to the further activities of this first cause, for we have left room for unending experiments in evolution These experiments must include Life The cosmologist can but set the stage for the biologist, and may rest content if he has found place for infinite variety of development We need suppose no interference with 'nature' later than creation But the inauguration of an infinity of experiments is something vastly transcending the Mosaic cosmology, and in an infinite multiplicity of Gardens of Eden the biologist need not feel that his activities are limited by bounds set by the cosmologist

155. One can say if one pleases that we have found God in the universe For the universe seems to be a perfect expression of those extra-temporal, extra-spatial attributes we should like to associate with the nature of God The universe, though describable in the space and time of an observer, is neither *in space* nor *in time*, to assert either would be meaningless There is no fundamental entity

'space', existing in itself, in which the world is placed, nor is there any fundamental stream 'time' in which the world is placed. Out of an observer's temporal experience we have constructed observed time, and out of the same time-measures observer's space. But there is no underlying time or space with which these can conceivably be correlated. Thus the universe, and with it, if we posit one, its Creator, is neither in space nor in time. The event of creation can be placed prior to the experiences of any individual. But no meaning could be attached to asking what was prior to creation. For as there were no observers to experience a temporal sequence, the notion of a temporal experience, and so of time, prior to creation, is without any significance.

156. Nevertheless though the attributes some would wish to associate with deity are found in the universe, it seems difficult to identify God with the universe, owing to the once-given-ness of the universe, or creation antecedent to experience. This is an ancient topic. Our contribution here is merely to point out that our analysis of the universe though giving the answer, and a substantial answer, to the question of its 'how', 'where', and 'when', gives no answer to its 'why'. It is by no means certain that we are entitled to ask the question 'Why the universe?' But those who feel the question to be a permissible one can legitimately answer the question 'why?' by positing God. The physicist and cosmologist then need God only once, to ensure creation. For the biologist, the world provides further opportunity for divine planning, if we admit the possibility of not entirely coincident evolutionary trends in similar circumstances. For man as more than cosmologist, as more than biologist, as possessing mind, possibly endowed with an immortal soul, God is perhaps needed always. Theoretical cosmology is but the starting-point for deeper philosophical inquiries.†

† The views here developed may appear somewhat out of place in a work devoted to the mathematical consequences of general notions derived from experience. But to omit a statement of the position and the questions and answers suggested by the mathematical development would be to flinch before important issues. I am aware of the dangers attendant upon an excursion into theology, but a statement of private conclusions may be admissible. In extenuation I would remark that the mathematical investigations which follow, like those which precede, require no assent to the non-mathematical portions of this chapter.

PART III

THE CAREER OF THE UNIVERSE

VIII

WORLD-TRAJECTORIES

157. We now resume our mathematical investigations at the point at which we left off at the end of Chapter V

We are now to discuss the trajectory of a *free* particle in the presence of the distribution of *given* particles in motion we have called the simple kinematic world-model That is to say, given the set of particles moving according to the law

$$\mathbf{V} = \mathbf{P}/t \tag{1}$$

and distributed according to the law

$$n\,dxdydz = \frac{Bt\,dxdydz}{c^3(t^2-\mathbf{P}^2/c^2)^2}, \tag{2}$$

all reckoned in the space- and time-coordinates of the observer at the origin, we desire to discuss how this observer will describe the motion of an additional free particle projected in any manner We have already seen that the equations governing the free motion are

$$\frac{d\mathbf{P}}{dt} = \mathbf{V}, \tag{3}$$

$$\frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi), \tag{4}$$

where

$$\left.\begin{aligned} X = t^2-\frac{\mathbf{P}^2}{c^2}, \qquad Y = 1-\frac{\mathbf{V}^2}{c^2}, \qquad Z = t-\frac{\mathbf{P}\ \mathbf{V}}{c^2}, \\ \xi = Z^2/XY \end{aligned}\right\} \tag{5}$$

We shall now show that, without any knowledge of the precise form of the function $G(\xi)$, these equations may be integrated completely, and the general nature of the trajectories ascertained In succeeding chapters we shall ascertain the function $G(\xi)$ more precisely, and examine the properties of systems of trajectories of free particles We have already discussed the particular set of trajectories $\mathbf{P} = \mathbf{V}_0 t$, $\mathbf{V} = \mathbf{V}_0$, which are in fact the trajectories of the *given* particles

Constants of integration

158. We first notice that the integrals may be expected to contain 6 arbitrary constants of integration For a particle may be projected with an arbitrary velocity **V** from an arbitrary position **P** at any given epoch t Given any set of initial values of $\mathbf{P}, \mathbf{V}, t$, we shall expect to be able to trace not only the forward but also the backward motion of the particle, i e not only its future trajectory but also the past trajectory it has already traversed if it be supposed to have arrived at P at epoch t with velocity **V** without undergoing any collisions This past history must extend back to $t = 0$ At $t = 0$ the given system reduces to the singularity $\mathbf{P} = 0$ $(x = 0, y = 0, z = 0)$ Hence we shall *expect* that the trajectory passes through† $t = 0$, $\mathbf{P} = 0$ Hence all trajectories pass through this point, whatever their circumstances of projection, i e whatever the values of the 6 arbitrary constants defining the particular trajectory considered Three of these constants may be taken to be the three components of the vector velocity $\mathbf{V}_0$ which the free particle possessed at $t = 0$, $\mathbf{P} = 0$ It follows that $t = 0$, $\mathbf{P} = 0$ is a singular point on all trajectories, for at this point the 'initial' conditions at $t = 0$, $\mathbf{P} = 0$, $\mathbf{V} = \mathbf{V}_0$, do not serve fully to define the trajectory A triple infinity of trajectories, corresponding to the 3 other arbitrary constants, will pass through $t = 0$, $\mathbf{P} = 0$ with a given velocity $\mathbf{V} = \mathbf{V}_0$ We shall later show that such a triple infinity constitutes a sub-system of trajectories physically associated with the *fundamental particle* $\mathbf{V} = \mathbf{V}_0$ of the original kinematic model, and we shall identify such a sub-system as defining the nebular sub-system moving with the *constant* velocity $\mathbf{V} = \mathbf{V}_0$

In this chapter we consider only a single free particle, and ascertain the physical significance of the three remaining constants of integration other than those prescribed by the value of $\mathbf{V}_0$ We shall obtain in succession what may be called the 'X-integral', the 'vector-integrals', and the 'Λ-integral '

Preliminary lemma

159. It is readily proved that $\xi > 1$, i e $Z^2 > XY$, unless $\mathbf{P} = \mathbf{V}t$, in which case $\xi = 1$, $Z^2 = XY$ For

$$Z^2 > XY$$

† This will be shown later analytically

when $X > 0$, $Y > 0$, provided

$$\left(t-\frac{\mathbf{P}\cdot\mathbf{V}}{c^2}\right)^2 > \left(t^2-\frac{\mathbf{P}^2}{c^2}\right)\left(1-\frac{\mathbf{V}^2}{c^2}\right),$$

i e provided

$$\mathbf{V}^2t^2-2(\mathbf{P}\cdot\mathbf{V})t+\mathbf{P}^2\left(1-\frac{\mathbf{V}^2}{c^2}\right)+\frac{(\mathbf{P}\cdot\mathbf{V})^2}{c^2} > 0 \tag{6}$$

This inequality holds good provided

$$(\mathbf{P}\cdot\mathbf{V})^2 < \mathbf{V}^2\left[\mathbf{P}^2\left(1-\frac{\mathbf{V}^2}{c^2}\right)+\frac{(\mathbf{P}\cdot\mathbf{V})^2}{c^2}\right],$$

i e since $1-\mathbf{V}^2/c^2 > 0$, provided

$$(\mathbf{P}\cdot\mathbf{V})^2 < \mathbf{P}^2\mathbf{V}^2$$

or†

$$(\mathbf{P}\wedge\mathbf{V})^2 > 0$$

This is satisfied unless $\mathbf{P}$ is parallel to $\mathbf{V}$ If $\mathbf{P}\wedge\mathbf{V} = 0$, the inequality becomes, by (6),

$$(\mathbf{P}-\mathbf{V}t)^2 > 0,$$

which is again satisfied unless $\mathbf{P} = \mathbf{V}t$ If $\mathbf{P} = \mathbf{V}t$, we have $Z^2 = XY$, or $\xi = 1$

The X-integral

160 By straightforward differentiation of (5) with respect to the time t and use of (3) and (4) we find

$$\frac{dX}{dt} = 2Z, \quad \frac{dY}{dt} = -2\frac{Y}{X}(tY-Z)G(\xi), \quad \frac{dZ}{dt} = Y\left[1+G(\xi)\frac{X-tZ}{X}\right], \tag{7}$$

whence

$$\frac{d\xi}{dt} = -\frac{2Z}{X}(\xi-1)(1+G(\xi)) \tag{7'}$$

Hence

$$\frac{1}{X}\frac{dX}{d\xi} = -\frac{1}{(\xi-1)(1+G(\xi))} \tag{8}$$

The integral of this is

$$X = X_1\exp\left[-\int\limits_{\xi_1}^{\xi}\frac{d\xi}{(\xi-1)(1+G(\xi))}\right], \tag{9}$$

where X_1 is the value taken by X at the point of the trajectory at which ξ takes the arbitrary value ξ_1, X_1 and ξ_1 together amount to a single arbitrary constant of integration Later we shall be able to

† The expression $\mathbf{P}\wedge\mathbf{V}$ denotes the vector products of $\mathbf{P}$ and $\mathbf{V}$ in this order The identity $(\mathbf{P}\wedge\mathbf{V})^2 = \mathbf{P}^2\mathbf{V}^2-(\mathbf{P}\cdot\mathbf{V})^2$ is well known It is equivalent to Lagrange's identity in algebra, or can be proved directly by vector methods

take $\xi_1 = \infty$, but for the present we retain greater flexibility and control by keeping ξ_1 arbitrary

161. By (9), X has a constant sign, namely the sign of X_1 If at any point of the trajectory the particle P lies inside the expanding sphere $|\mathbf{P}| = ct$, then at this point $X \equiv t^2 - \mathbf{P}^2/c^2$ is positive, and so X is positive along the whole trajectory Hence along the whole trajectory $|\mathbf{P}| < ct$, and so the trajectory lies entirely *inside* the expanding sphere $|\mathbf{P}| = ct$

162. Formally, the equations of motion (3) and (4) may be supposed to hold good *outside* the sphere $|\mathbf{P}| = ct$, for which events X is negative, and ξ is negative But if X is once negative, it is always negative along the same trajectory, and so such a trajectory lies wholly *outside* the expanding sphere $|\mathbf{P}| = ct$ The expanding light-sphere $|\mathbf{P}| = ct$ is thus an impassable frontier separating trajectories into two families, no particle ever crosses this frontier Further, since the behaviour of particles inside the expanding sphere depends entirely on the values of $G(\xi)$ for $1 \leqslant \xi < \infty$, whilst the behaviour of particles outside would depend on the values of $G(\xi)$ for $\xi < 0$, there is no interaction whatever between phenomena outside the sphere and phenomena inside it, if we like to make use of a theory of causation, we should say that no causal influence whatever can cross the frontier $|\mathbf{P}| = ct$ We have already seen that in any case phenomena outside $|\mathbf{P}| = ct$ would be completely unobservable, since to an observer inside $|\mathbf{P}| = ct$ the sky is 'everywhere dense' in fundamental particles lying inside $|\mathbf{P}| = ct$, so that there are no 'holes' or 'windows' through which observation could be made, or light pass The mathematical analysis is thus completely self-consistent Absence of the passage of causal influence across $|\mathbf{P}| = ct$ is identical with the complete impossibility of making observations on events on the other side of $|\mathbf{P}| = ct$ We can perfectly well describe and analyse conceptual trajectories outside $|\mathbf{P}| = ct$, we can perfectly well fill the 'space' outside $|\mathbf{P}| = ct$ with a system of particles in motion satisfying the cosmological principle, as we shall see later But to do so is entirely irrelevant to any phenomena observable by an observer inside $|\mathbf{P}| = ct$ Since nothing is gained by contemplating experiences outside $|\mathbf{P}| = ct$ we shall henceforward not consider them Thus always $X_1 > 0$ and $X \geqslant 0$ along all trajectories

163 Now trace any particular trajectory backwards as $t \to 0$ Then since $X \equiv t^2 - \mathbf{P}^2/c^2 \geqslant 0$, it follows that, as $t \to 0$, $|\mathbf{P}| \to 0$ along the trajectory Thus the trajectory originates from the event $t = 0$, $\mathbf{P} = 0$ At this event, $X = 0$ If the particle passes through the point $\mathbf{P} = 0$ at $t = 0$ with velocity $\mathbf{V}_0$, then, near $t = 0$, $\mathbf{V} \sim \mathbf{V}_0$, $\mathbf{P} \sim \mathbf{V}_0 t$, and $\xi = [t - (\mathbf{P}\ \mathbf{V})/c^2]^2/(t^2 - \mathbf{P}^2/c^2)(1 - \mathbf{V}^2/c^2) \sim 1$ Also as $t \to 0$, and $X \to 0$, $\xi \to 1$ Hence (9) demands that, as $\xi \to 1$,

$$-\int\limits_{\xi_1}^{\xi} \frac{d\xi}{(\xi-1)(1+G(\xi))} \to -\infty$$

Hence

$$\int\limits_{\xi \to 1} \frac{d\xi}{(\xi-1)(1+G(\xi))}$$

diverges, and to $-\infty$ We have already seen that $G(1)$ is a definite number defining the accelerations of resting or slowly moving particles reckoned with respect to a neighbouring fundamental particle The two results are consistent

Approximately, when $\xi \sim 1$,

$$X \sim X_1 e^{-\frac{\log(\xi-1)}{1+G(1)}} \times \text{const}$$

or

$$X \sim \frac{\text{const}}{(\xi-1)^{[1+G(1)]^{-1}}}$$

It follows, since $X \to 0$ as $\xi \to 1$, that $1+G(1)$ is negative, and therefore that $G(1)$ is negative This gives the phenomenon of *attraction* for a resting or slowly moving particle in the vicinity of a fundamental particle, *towards* the fundamental particle We have already seen that comparison with Newtonian dynamics and gravitation gave

$$-G(1) = \tfrac{4}{3}\pi \frac{\gamma}{t} \frac{mB}{c^3},$$

and this now shows that γ is positive We thus establish for the case of our system of fundamental particles that gravitation appears locally as a phenomenon of *attraction*, without recourse to any appeal to experience This is a fundamentally important result, and its establishment on kinematic grounds alone may be considered as an achievement of the present method of approach to gravitation

164 An exceptional case occurs if $|\mathbf{V}_0| = c$ In that case $Y = 0$, and the acceleration is zero The particle continues to move with the

velocity of light, and remains always on the sphere $|P| = ct$, and therefore unobservable. Such particles do not enter into consideration

165. It will prove convenient later to adopt the three constants defining the vector velocity $\mathbf{V}_0$ which the particle possessed at $t = 0$ as three of the integration constants of the trajectory Actually we shall not do so at present, but wait to see which integration constants, additional to X_1 make their appearance as we carry out further integrations We now seek five further integrals, corresponding to five further constants of integration Four of these appear without difficulty in the form of four 'vector integrals', which we now proceed to obtain

The vector integrals

166. From (3) and (4), we find at once

$$\frac{d}{dt}(\mathbf{P}-\mathbf{V}t) = -(\mathbf{P}-\mathbf{V}t)t\frac{Y}{X}G(\xi), \tag{10}$$

$$\frac{d}{dt}(\mathbf{P}\wedge\mathbf{V}) = -(\mathbf{P}\wedge\mathbf{V})t\frac{Y}{X}G(\xi) \tag{11}$$

Further, we find that

$$\frac{d}{dt}(Z^2-XY)^{\frac{1}{2}} = -(Z^2-XY)^{\frac{1}{2}}t\frac{Y}{X}G(\xi) \tag{12}$$

The details of the derivation of these are supplied in Note 4 It follows that (10) and (11) can be integrated in the form

$$\mathbf{P}-\mathbf{V}t = \mathbf{f}\theta, \tag{13}$$

$$\mathbf{P}\wedge\mathbf{V} = c\mathbf{l}\theta, \tag{14}$$

where $$\theta = +(Z^2-XY)^{\frac{1}{2}} \tag{15}$$

and $\mathbf{f}, \mathbf{l}$ are constant arbitrary vectors, (13) and (14) are accordingly integrals of the original equations of motion

167. The vectors $\mathbf{f}, \mathbf{l}$ are not, however, independent For we have identically

$$(\mathbf{P}-\mathbf{V}t)\ (\mathbf{P}\wedge\mathbf{V}) \equiv 0 \tag{16}$$

and $$\theta^2 \equiv \frac{1}{c^2}(\mathbf{P}-\mathbf{V}t)^2-\frac{1}{c^4}(\mathbf{P}\wedge\mathbf{V})^2 \tag{17}$$

It follows that $$\mathbf{f}\ \mathbf{l} = 0, \tag{18}$$

and, provided $\theta \neq 0$, that also

$$\mathbf{f}^2-\mathbf{l}^2 = c^2. \tag{19}$$

(If $\theta = 0$, then $\xi = 1$, and the trajectories are simply $\mathbf{P} = \mathbf{V}t$,

$\mathbf{V} = \text{const}$, which are the trajectories of the given fundamental particles For a freely projected particle not coinciding in velocity with one of the fundamental particles, necessarily $\theta \neq 0$)

It follows that since (18) and (19) are two scalar equations, the vectors $\mathbf{f}, \mathbf{l}$ correspond to six scalar constants connected by two relations, i e that $\mathbf{f}$ and $\mathbf{l}$ correspond to 4 independent constants of integration Together with X_1, they provide 5 of the desired 6 constants of integration

168. Equations (13) and (14) may be solved for $\mathbf{V}$ and $\mathbf{P}$ in the form (see Note 4)

$$\mathbf{V} = c\frac{\mathbf{l}\wedge\mathbf{f}}{\mathbf{f}^2}+\lambda\mathbf{f}, \tag{20}$$

$$\mathbf{P} = c\frac{\mathbf{l}\wedge\mathbf{f}}{\mathbf{f}^2}t+(\lambda t+\theta)\mathbf{f}, \tag{21}$$

where λ is a function of t not determined by the 5 integrals so far found

The Λ-integral

169. After I had obtained the above integrals, whose existence is obvious, it took me over twelve months to obtain the sixth integral This may be found simply as follows

Inserting (20) and (21) in the equation of motion (4) we have at once

$$\frac{d\lambda}{dt} = \theta\frac{Y}{X}G(\xi)$$

But by (7')

$$\frac{d\xi}{dt} = -\frac{2Z}{X}(\xi-1)(1+G(\xi))$$

Hence

$$\frac{d\lambda}{d\xi} = -\frac{1}{2}\frac{Y\theta}{Z}\frac{G(\xi)}{(\xi-1)(1+G(\xi))} \tag{22}$$

But

$$\frac{\theta}{Z} = \frac{(Z^2-XY)^{\frac{1}{2}}}{Z} = \frac{X^{\frac{1}{2}}Y^{\frac{1}{2}}}{Z}(\xi-1)^{\frac{1}{2}} = \frac{(\xi-1)^{\frac{1}{2}}}{\xi^{\frac{1}{2}}}, \tag{23}$$

and, from (20),

$$Y = 1-\frac{\mathbf{V}^2}{c^2} = 1-\frac{\mathbf{l}^2\mathbf{f}^2-(\mathbf{l}\ \mathbf{f})^2}{\mathbf{f}^4}-\frac{\lambda^2\mathbf{f}^2}{c^2},$$

or by (18) and (19)

$$Y = \frac{c^2}{\mathbf{f}^2}-\frac{\lambda^2\mathbf{f}^2}{c^2}.$$

Put

$$\lambda = \Lambda(c^2/\mathbf{f}^2), \tag{24}$$

where Λ is some new function of position along the trajectory Then

$$Y = \frac{c^2}{\mathbf{f}^2}(1-\Lambda^2) \tag{25}$$

Inserting (23), (24), and (25), in (22) we have

$$\frac{1}{1-\Lambda^2}\frac{d\Lambda}{d\xi} = -\frac{1}{2}\frac{G(\xi)}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1+G(\xi))}$$

$$= -\frac{1}{2}\left[\frac{1}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}} - \frac{1}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1+G(\xi))}\right]$$

This integrates in the form

$$\frac{1+\Lambda}{1-\Lambda} = \text{const}\ \frac{1}{[\xi^{\frac{1}{2}}+(\xi-1)^{\frac{1}{2}}]^2}\exp\left(\int^{\xi}\frac{d\xi}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1+G(\xi))}\right) \tag{26}$$

Now by (25), $\Lambda^2 < 1$, the value $\Lambda^2 = 1$ occurring only when $|\mathbf{V}| = c$ Hence if Λ_0 is the value of Λ at $t = 0$, when $\mathbf{V} = \mathbf{V}_0$, ($|\mathbf{V}_0| < c$), we have $\Lambda_0^2 < 1$ Hence as $t \to 0$, and $\xi \to 1$, the left-hand side of (26) tends to the finite non-zero positive number $(1+\Lambda_0)/(1-\Lambda_0)$ Hence the integral

$$\int^{\xi}\frac{d\xi}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1+G(\xi))}$$

is convergent at $\xi = 1$ Equation (26) may accordingly be written in the form

$$\frac{1+\Lambda}{1-\Lambda} = \frac{1+\Lambda_0}{1-\Lambda_0}\,\frac{\exp\left(\int_1^{\xi}\frac{d\xi}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1+G(\xi))}\right)}{[\xi^{\frac{1}{2}}+(\xi-1)^{\frac{1}{2}}]^2} \tag{27}$$

The constant Λ_0 may be considered as a new constant of integration, making with X_1 and the four equivalent to $\mathbf{f}$ and $\mathbf{l}$ six independent constants of integration, and Λ is a function of ξ and Λ_0 only The variable θ is given in terms of ξ and the constants of integration by

$$\theta = X^{\frac{1}{2}}Y^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}} = \frac{cX^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1-\Lambda^2)^{\frac{1}{2}}}{|\mathbf{f}|} \tag{28}$$

Equations (20) and (21) may now be written

$$\mathbf{V} = c\frac{\mathbf{l}\wedge\mathbf{f}}{\mathbf{f}^2}+\Lambda\frac{c^2}{\mathbf{f}^2}\mathbf{f}, \tag{29}$$

$$\mathbf{P} = c\frac{\mathbf{l}\wedge\mathbf{f}}{\mathbf{f}^2}t+\theta\mathbf{f}+\Lambda t\frac{c^2}{\mathbf{f}^2}\mathbf{f} \tag{30}$$

These express $\mathbf{P}$ and $\mathbf{V}$ as functions of ξ and t and the six constants of integration It remains to determine t This can be done without further quadrature

The time-integral

170 We make use of the identity

$$Z \equiv X^{\frac{1}{2}}Y^{\frac{1}{2}}\xi^{\frac{1}{2}} = t - \frac{\mathbf{P}\ \mathbf{V}}{c^2},$$

which by (25) may be written

$$t - \frac{\mathbf{P}\ \mathbf{V}}{c^2} = X^{\frac{1}{2}}\xi^{\frac{1}{2}}\frac{c}{|\mathbf{f}|}(1-\Lambda^2)^{\frac{1}{2}} \tag{31}$$

Now by scalar multiplication of (29) and (30) we have

$$\frac{\mathbf{P}\ \mathbf{V}}{c^2} = \frac{\mathbf{l}^2}{\mathbf{f}^2}t + \Lambda\theta + \Lambda^2 t\frac{c^2}{\mathbf{f}^2},$$

or using (19),

$$\frac{\mathbf{P}\ \mathbf{V}}{c^2} = \frac{\mathbf{f}^2 - c^2}{\mathbf{f}^2}t + \Lambda\theta + \Lambda^2 t\frac{c^2}{\mathbf{f}^2}$$

Inserting this in (31), we have

$$t\frac{c^2}{\mathbf{f}^2}(1-\Lambda^2) = \Lambda\theta + X^{\frac{1}{2}}\xi^{\frac{1}{2}}\frac{c}{|\mathbf{f}|}(1-\Lambda^2)^{\frac{1}{2}}$$

Substituting in this for θ from (28), we have

$$t\frac{c^2}{\mathbf{f}^2}(1-\Lambda^2) = \frac{c}{|\mathbf{f}|}X^{\frac{1}{2}}(1-\Lambda^2)^{\frac{1}{2}}[\xi^{\frac{1}{2}} + \Lambda(\xi-1)^{\frac{1}{2}}],$$

or

$$t = \frac{|\mathbf{f}|}{c}\frac{X^{\frac{1}{2}}}{(1-\Lambda^2)^{\frac{1}{2}}}[\xi^{\frac{1}{2}} + \Lambda(\xi-1)^{\frac{1}{2}}] \tag{32}$$

171 Equation (32) determines t as a function of ξ and the 6 constants of integration Use of this in (29) and (30) then determines $\mathbf{V}$ and $\mathbf{P}$ as functions of ξ and the 6 constants of integration The resulting system of equations determines the march of $\mathbf{V}, \mathbf{P}, t$ along the trajectory fixed by the 6 constants of integration included in $X_1, \mathbf{f}, \mathbf{l}, \Lambda_0$ in terms of a parameter ξ which always satisfies identically the relation

$$\xi \equiv \frac{[t-(\mathbf{P}\ \mathbf{V})/c^2]^2}{(t^2-\mathbf{P}^2/c^2)(1-\mathbf{V}^2/c^2)}$$

The integration problem is thus completely solved

Use of unit vectors

172 The equations so found are capable of enormous simplification Instead of seeking such simplifications by analytical methods we attempt to gain insight into the physical meaning of the equations we have found, and use the physical results to suggest new analytical forms

Take three mutually perpendicular unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ forming a positive triad, i e such that

$$\mathbf{i}^2 = 1, \quad \mathbf{j}\,\mathbf{k} = 0, \quad \mathbf{j}\wedge\mathbf{k} = \mathbf{i},$$

etc , and put

$$\mathbf{f} = f\mathbf{i},$$

$$\mathbf{l} = l\mathbf{k} = (f^2-c^2)^{\frac{1}{2}}\mathbf{k},$$

so that

$$\mathbf{l}\wedge\mathbf{f} = f(f^2-c^2)^{\frac{1}{2}}\mathbf{j}.$$

Then by (28), (29), (30), and (32) we have

$$\mathbf{V} = \frac{c(f^2-c^2)^{\frac{1}{2}}}{f}\mathbf{j}+\frac{c^2}{f}\Lambda\mathbf{i}, \tag{33}$$

$$\mathbf{P} = \frac{c(f^2-c^2)}{f}\mathbf{j}t+\frac{c^2}{f}\Lambda\mathbf{i}t+cX^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1-\Lambda^2)^{\frac{1}{2}}\mathbf{i}, \tag{34}$$

$$t = \frac{f}{c}\frac{X^{\frac{1}{2}}}{(1-\Lambda^2)^{\frac{1}{2}}}[\xi^{\frac{1}{2}}+\Lambda(\xi-1)^{\frac{1}{2}}], \tag{35}$$

Λ being a function of ξ defined by (27) These equations express $\mathbf{P}, \mathbf{V}, t$ parametrically in terms of the variable ξ and 6 constants of integration $X_1, f, \Lambda_0, \mathbf{i}$ (equivalent to 2 constants) and $\mathbf{j}$ (equivalent to one constant) The positive values of all surds are implied

173. From the equation preceding (22), it follows that $1+G(\xi)$ cannot vanish for finite ξ unless $d\xi/dt$ vanishes This would mean $dt/d\xi = \infty$, which is physically impossible along a trajectory, at a finite point of it passed through with a velocity less than c Hence $1+G(\xi)$ never vanishes so long as $\xi < \infty$ † But $1+G(1)$ is negative Hence $1+G(\xi)$ is always negative Hence the integral in the exponential factor in (27) is always negative Hence as ξ increases to infinity, this exponential either tends to a finite limit or tends to zero, also the denominator in (27) tends to infinity as $\xi \to \infty$ Hence as ξ increases to infinity the right-hand side of (27) tends to zero Hence as $\xi \to \infty$, $\Lambda \to -1$, and, by (25), $|\mathbf{V}| \to c$ Moreover, $1+G(\xi)$ and $G(\xi)$ are always both negative, and so, by the equation following (25), $d\Lambda/d\xi$ is constantly negative and Λ steadily decreases as ξ increases Thus as ξ passes from 1 to infinity, Λ decreases from Λ_0 to -1

Hodograph of the motion

174. The formulae for $\mathbf{P}$ and $\mathbf{V}$ show that the trajectory lies in the plane of $\mathbf{i}$ and $\mathbf{j}$, and thus lies in a plane passing through the origin—

† I shall use the notation $\xi < \infty$ to imply that ξ has not yet passed through ∞, and the notation $\infty > \xi$ to imply that ξ has passed through ∞

which, be it remembered, is any arbitrary particle of the fundamental system We can now construct the *hodograph* of the motion as follows.

Choose an arbitrary particle O of the given fundamental set $\mathbf{V} = \text{const}$, $\mathbf{P} = \mathbf{V}t$ as origin, and regard it as at rest Take a unit vector $\mathbf{j}$ in an arbitrary direction through O and mark off a length $OH = \dfrac{c(f^2-c^2)^{\frac{1}{2}}}{f}\mathbf{j}$ along it, where f ($> c$) is arbitrary Through H, in an arbitrary azimuth draw a unit vector $\mathbf{i}$ perpendicular to $\mathbf{j}$ and mark off the length $HM_0 = (c^2/f)\Lambda_0\,\mathbf{i}$ along it, where Λ_0 is arbitrary and may be positive or negative subject to $|\Lambda_0| < 1$ Further mark off $HM_l = -(c^2/f)\mathbf{i}$

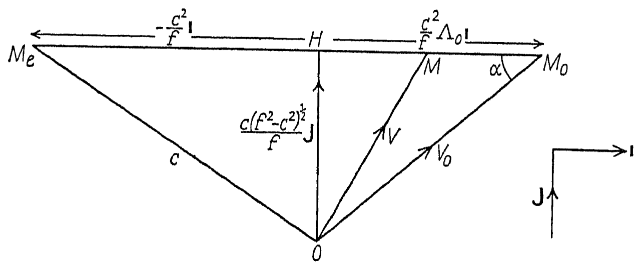


FIG 13 Hodograph of a free particle

Then if M is any point between M_0 and M_l, by (33), OM represents the velocity of the particle P for the position on its trajectory at which $HM = (c^2/f)\Lambda(\xi)$ OM_0 represents the velocity of P at $t = 0$, and the length of OM_l is c As t advances from $t = 0$, the representative point M moves from M_0 along $M_0 H$, and tends to M_l as $\xi \to \infty$ The position M_l corresponds to $|\mathbf{V}| = c$

175. The reader will naturally ask whether M arrives at M_l in a finite time or only as $t \to \infty$ Since as $\xi \to \infty$, $\Lambda \to -1$, formula (35) for t tends to assume the form $0/0$ We shall find later (to my own personal astonishment) that M arrives at M_l in a *finite* time, then retraces its path along the hodograph locus, finally coming to rest at a determinate point M_0' on $M_l M_0$ This means that along the trajectory the particle P attains the velocity of light at a finite epoch, then decelerates and ultimately assumes a constant limiting velocity as $t \to \infty$ The physical and astronomical consequences of this behaviour are of the first importance But to handle the expressions we have obtained for $\mathbf{P}, \mathbf{V}, t$ it is convenient now to choose new constants of integration which are functions of the old ones

Choice of new constants of integration

176 We have seen that $OM_0 = \mathbf{V}_0$ the vector velocity at $t = 0$, we write $V_0 = |\mathbf{V}_0|$ Let α $(0 \leqslant \alpha \leqslant \pi)$ be the angle between $\mathbf{V}_0$ and $\mathbf{i}$ Then $\mathbf{V}_0 \cdot \mathbf{i} = V_0 \cos\alpha$, and so from the diagram

$$\frac{c^2}{f}\Lambda_0 = V_0 \cos\alpha,$$

$$\frac{c(f^2-c^2)^{\frac{1}{2}}}{f} = V_0 \sin\alpha,$$

whence

$$f = \frac{c}{[1-(V_0^2 \sin^2\alpha)/c^2]^{\frac{1}{2}}}, \tag{36}$$

$$\Lambda_0 = \frac{(V_0 \cos\alpha)/c}{[1-(V_0^2 \sin^2\alpha)/c^2]^{\frac{1}{2}}}, \tag{37}$$

$$1-\Lambda_0^2 = \frac{1-V_0^2/c^2}{1-(V_0^2 \sin^2\alpha)/c^2} \tag{38}$$

By (33), we have

$$\mathbf{V}_0 = \frac{c(f^2-c^2)^{\frac{1}{2}}}{f}\mathbf{j} + \frac{c^2}{f}\Lambda_0 \mathbf{i}, \tag{39}$$

whence

$$\mathbf{V} = \mathbf{V}_0 + \frac{c^2}{f}(\Lambda-\Lambda_0)\mathbf{i}, \tag{40}$$

and

$$\mathbf{P} = \mathbf{V}_0 t + \frac{c^2}{f}(\Lambda-\Lambda_0)t\mathbf{i} + cX^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1-\Lambda^2)^{\frac{1}{2}}\mathbf{i} \tag{41}$$

177. These formulae suggest the following substitutions Introduce a parameter ϵ, a function of $\mathbf{V}_0$ and $\mathbf{i}$, defined by

$$\sinh\epsilon = \frac{(V_0 \cos\alpha)/c}{(1-V_0^2/c^2)^{\frac{1}{2}}}, \qquad \cosh\epsilon = \frac{[1-(V_0^2 \sin^2\alpha)/c^2]^{\frac{1}{2}}}{(1-V_0^2/c^2)^{\frac{1}{2}}} \tag{42}$$

Then

$$f = \frac{c \operatorname{sech}\epsilon}{(1-V_0^2/c^2)^{\frac{1}{2}}}, \tag{43}$$

$$\Lambda_0 = \tanh\epsilon, \tag{44}$$

$$(1-\Lambda_0^2)^{\frac{1}{2}} = \operatorname{sech}\epsilon \tag{45}$$

Further, define a variable ζ in terms of ξ by

$$\cosh\zeta = +\xi^{\frac{1}{2}}, \qquad \sinh\zeta = +(\xi-1)^{\frac{1}{2}}, \tag{46}$$

so that

$$\xi^{\frac{1}{2}} + (\xi-1)^{\frac{1}{2}} = e^{\zeta},$$

and in formula (27) for Λ write

$$\eta = -\frac{1}{2}\int_1^{\xi} \frac{d\xi}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}[1+G(\xi)]} \tag{47}$$

Then by (27), $$\frac{1+\Lambda}{1-\Lambda} = \frac{1+\Lambda_0}{1-\Lambda_0}e^{-2(\eta+\zeta)} = e^{2(\epsilon-\eta-\zeta)},$$
so that
$$\Lambda = \tanh(\epsilon-\eta-\zeta), \tag{48}$$
$$(1-\Lambda^2)^{\frac{1}{2}} = \operatorname{sech}(\epsilon-\eta-\zeta), \tag{49}$$
$$\Lambda-\Lambda_0 = -\frac{\sinh(\eta+\zeta)}{\cosh\epsilon\cosh(\epsilon-\eta-\zeta)}. \tag{50}$$
We now regard $\mathbf{V}_0$ (3 constants), $\mathbf{i}$ (2 constants), and X_1 as our 6 constants of integration

The $\mathbf{P}, \mathbf{V}, t$ *integrals*

178 Insert from (42)–(50) for f, Λ_0, Λ, and ξ in (33), (34), (35) in turn Then after a little straightforward hyperbolic trigonometry we find
$$\mathbf{V} = \mathbf{V}_0 - c(1-V_0^2/c^2)^{\frac{1}{2}}\frac{\sinh(\eta+\zeta)}{\cosh(\epsilon-\eta-\zeta)}\mathbf{i}, \tag{51}$$
$$\mathbf{P} = \mathbf{V}_0 t - cX^{\frac{1}{2}}\frac{\sinh\eta}{\cosh\epsilon}\mathbf{i}, \tag{52}$$
$$t = \frac{X^{\frac{1}{2}}}{(1-V_0^2/c^2)^{\frac{1}{2}}}\frac{\cosh(\epsilon-\eta)}{\cosh\epsilon} \tag{53}$$
These are found to yield the useful formulae
$$\mathbf{P}-\mathbf{V}t = cX^{\frac{1}{2}}\frac{\sinh\zeta}{\cosh(\epsilon-\eta-\zeta)}\mathbf{i}, \tag{54}$$
$$Y^{\frac{1}{2}} = (1-V_0^2/c^2)^{\frac{1}{2}}\frac{\cosh\epsilon}{\cosh(\epsilon-\eta-\zeta)} \tag{55}$$
For convenience we repeat here the X-integral and η-integral
$$X = X_1\exp\left[-\int\limits_{\xi_1}^{\xi}\frac{d\xi}{(\xi-1)[1+G(\xi)]}\right], \tag{56}$$
$$\eta = -\frac{1}{2}\int\limits_{1}^{\xi}\frac{d\xi}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}[1+G(\xi)]}, \tag{57}$$
and the definitions of ϵ and ζ,
$$\sinh\epsilon = \frac{(\mathbf{V}_0\ \mathbf{i})/c}{(1-V_0^2/c^2)^{\frac{1}{2}}}, \tag{58}$$
$$\cosh\zeta = \xi^{\frac{1}{2}} \tag{59}$$
As t increases from 0, ξ increases from 1, ζ and η increase from 0

179 Relations (51), (52), (53) define the trajectory of any free particle in terms of a parameter ξ and the 6 arbitrary constants of integration defined by $\mathbf{V}_0$, $\mathbf{i}$, and X_1 they enumerate the totality of all possible trajectories or arcs of trajectories, other than the fundamental rectilinear trajectories $\mathbf{V} =$ const , $\mathbf{P} = \mathbf{V}t$

Properties of the $\mathbf{P}, \mathbf{V}, t$ *integrals*

180. I pause here to point out that (51), (52), (53) are very remarkable formulae

(1) They are invariant in form for a Lorentz transformation from the arbitrary particle-observer O taken as origin to any other particle-observer O' of the fundamental set $\mathbf{P} = \mathbf{U}t$, $\mathbf{U} =$ const , for arbitrary $\mathbf{U}$ The parameter ξ is itself invariant under such a transformation The constant X_1 is not only constant along any trajectory but invariant under the transformation from O to O' The constants $\mathbf{V}_0$ and $\mathbf{i}$ transform into new constants $\mathbf{V}_0'$ and $\mathbf{i}'$ (depending on $\mathbf{U}$) which can be found without trouble The reader will find it an interesting exercise to verify this invariance of form in detail

(2) They satisfy identically the equations of motion (3) and (4), for arbitrary $\mathbf{V}_0, \mathbf{i}, X_1$

(3) Their form is independent of the form of the function $G(\xi)$

(4) They satisfy identically the two relations

$$t^2 - \mathbf{P}^2/c^2 \equiv X,$$

$$\frac{[t-(\mathbf{P}\ \mathbf{V})/c^2]^2}{(t^2-\mathbf{P}^2/c^2)(1-\mathbf{V}^2/c^2)} \equiv \cosh^2\zeta \equiv \xi$$

Thus ξ may be considered either as a number given by the original kinematic definition $\xi = Z^2/XY$ or as a scalar parameter varying along the trajectory and providing the variable of integration in (56) and (57)

(5) There is always one particle of the fundamental set $\mathbf{P} = \mathbf{U}t$ which possessed at $t = 0$ the same velocity as the particle whose trajectory we are considering For this fundamental particle, $\mathbf{U} = \mathbf{V}_0$ † We are fully entitled to take this fundamental particle as origin O

† Of course, since the event $t = 0$, $P = 0$ is a singular point on every trajectory, the fixation of $\mathbf{V}_{\binom{t=0}{P=0}}$ as equal to $\mathbf{U}$ does not uniquely determine the trajectory, as we have seen

This is equivalent to taking $\mathbf{V}_0 = 0$, in which case $\epsilon = 0$. For $\mathbf{V}_0 = 0$, the trajectory becomes

$$\mathbf{V} = -c\tanh(\eta+\zeta)\mathbf{\imath}, \tag{51'}$$

$$\mathbf{P} = -cX^{\frac{1}{2}}\sinh\eta\,\mathbf{i}, \tag{52'}$$

$$t = X^{\frac{1}{2}}\cosh\eta, \tag{53'}$$

$$\mathbf{P}-\mathbf{V}t = cX^{\frac{1}{2}}\frac{\sinh\zeta}{\cosh(\eta+\zeta)}\mathbf{\imath}, \tag{54'}$$

$$Y^{\frac{1}{2}} = \operatorname{sech}(\eta+\zeta) \tag{55'}$$

Thus seen from the fundamental particle which moves with a constant velocity equal to the velocity of P at $t = 0$, the trajectory appears as a straight line described in the direction of the vector $-\mathbf{\imath}$ with non-uniform velocity

181 Reckoned from this origin, the acceleration is given, by (4), (54′), and (55′), by

$$\frac{d\mathbf{V}}{dt} = \frac{c}{X^{\frac{1}{2}}}\frac{\sinh\zeta}{\cosh^3(\eta+\zeta)}G(\xi)\mathbf{\imath}$$

Since $G(\xi)$ is negative for $\xi < \infty$ the acceleration is in the direction of the vector $-\mathbf{\imath}$, and so is in the direction of motion, *positive outwards* from O. This is a remarkable result. It must be carefully distinguished from our previous result that a particle near O, at rest relative to O, has an acceleration $+\mathbf{P}G(1)/t^2$ *directed towards* O. The reconciliation of these apparently contradictory results will engage our attention shortly. In the meantime it may be mentioned that the outward acceleration from O of a particle P associated with O in the sense that the 'initial' ($t = 0$) velocity of P is equal to the constant velocity of O is the first hint we have obtained as to the origin of the outward motions apparent in the observed *forms* of the spiral nebulae. It is indeed the first hint that has ever been obtained on this topic in the history of cosmology. A free particle moving 'initially' with the same velocity as the nucleus of a nebula, and therefore tending in the first instance to move in the company of the nucleus, is accelerated outwards from the nucleus in the experience of an observer at the nucleus. We shall later show how the sub-system of freely moving particles which constitutes a nebula is built up out of such trajectories. For variation of $\mathbf{i}$ and X_1, for an assigned value of $\mathbf{V}_0$, there will be a triple infinity of such trajectories possible, corresponding to a possible triple infinity of particles forming a nebular sub-system, out

of the sextuple infinity of trajectories forming the grand world-system of nebular sub-systems

182. The discussion of the grand system of nebular sub-systems requires a far more delicate investigation than the motion of a single free particle For each free particle is then to be considered as moving not only in the presence of the given fundamental particles, but as also in the presence of the sextuple infinity of other free particles, and our analysis will require reconsideration from the start Before we embark on such an investigation, we shall proceed with our examination of the motion of a single free particle, and obtain the geometrical meaning of formulae (51), (52), (53) for the integrated trajectory We remind the reader at this stage that though we are considering the complete trajectory of any particle from $t = 0$ on, unimpeded by collisions, the investigation applies equally to any arc which is a portion of the trajectory For any free particle, however projected, from the circumstances $(\mathbf{P}, \mathbf{V})$ of its projection at time t, we can always calculate the trajectory constants $\mathbf{V}_0, \mathbf{i}, X_1$ by reversal of formulae (51), (52), (53)

Geometry of trajectories

183 Consider as before the hodograph of any one trajectory This hodograph depends only on the constants $\mathbf{V}_0$ and $\mathbf{i}$, and is independent of the constant X_1 Thus the single infinity of trajectories with the same $\mathbf{V}_0$ and $\mathbf{i}$ and different X_1's possesses the same hodograph

This hodograph is the line $M_0 M_l$, in the direction of $-\mathbf{i}$, where $\mathbf{i}$ is an arbitrary vector and OM_0 is the 'initial' velocity $\mathbf{V}_0$, O being the position of the observer making the calculations At the epoch t in the experience of O, let OM_0 be produced to P_0, where $OP_0 = \mathbf{V}_0 t$ The vector OP_0 in Fig 14 represents $\mathbf{V}_0 t$ on the same scale as OM_0 represents $\mathbf{V}_0$, so that $|OP_0|\ |OM_0| = t\ 1$ Draw a transversal through P_0 parallel to $-\mathbf{i}$, and so in the plane defined by OM_0 and $\mathbf{i}$ Then by (53), at the epoch t in O's experience, the particle P, projected from O at $t = 0$ with velocity $\mathbf{V}_0$, is at some point P on the transversal $-\mathbf{i}$ through P_0 For $P_0 P = OP - OP_0 = \mathbf{P} - \mathbf{V}_0 t =$ a vector parallel to $-\mathbf{i}$ If M, on the hodograph $M_0 M_l$ (the $-\mathbf{i}$ transversal through M_0), is the point which represents the velocity-vector $\mathbf{V}$ at the same epoch t, then OM represents $\mathbf{V}$, the velocity at P, and so the tangent to the trajectory at P is parallel to OM The trajectory is

tangent to OM_0 at O, and its general form is indicated by the dotted curved line The actual curvature of the dotted line depends on the form of $G(\xi)$, as also does the position of P on the transversal $-\mathbf{i}$ through P_0

Let OM meet this transversal in P_c, then OP_c, by similar triangles OM_0M, OP_0P_c, represents the vector $\mathbf{V}t$ Accordingly P_cP is the vector $\mathbf{P}-\mathbf{V}t$, and the acceleration at P, being proportional to

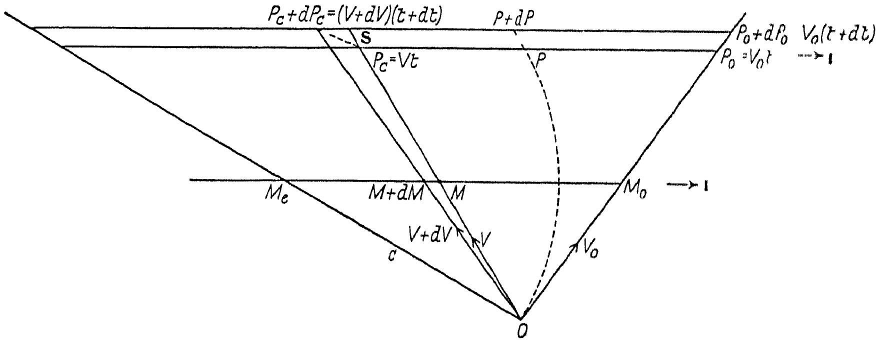


FIG 14 Geometry of free trajectories

$(\mathbf{P}-\mathbf{V}t)G(\xi)$, is parallel and opposite to P_cP, since $G(\xi)$ is negative P_c is in fact the position, at epoch t, of that fundamental particle which has been constantly moving, since $t=0$, with a velocity equal to the velocity $\mathbf{V}$ instantaneously possessed by the free particle at P, all in the reckoning of O But P is at rest, instantaneously, relative to the fundamental particle at P_c, and so at rest in the reckoning of the particle-observer moving with P_c But in the reckoning of this particle-observer, P_c is the centre of symmetry of the system of fundamental particles Thus the point P_c, obtained by the above construction, is the centre of the system in the frame in which P is momentarily at rest, and the acceleration of P is directed towards P_c in the reckoning of particle-observer O, and equally in the reckoning of the particle-observer P_c

184 We now prove that P lies between P_0 and P_c For

$$P_cP_0 = OP_0-OP_c = (\mathbf{V}_0-\mathbf{V})t = cX^{\frac{1}{2}}\frac{\sinh(\eta+\zeta)\cosh(\epsilon-\eta)}{\cosh(\epsilon-\eta-\zeta)\cosh\epsilon}\mathbf{i}$$

and $$PP_0 = OP_0-OP = \mathbf{V}_0t-\mathbf{P} = cX^{\frac{1}{2}}\frac{\sinh\eta}{\cosh\epsilon}\mathbf{i}$$

But elementary hyperbolic trigonometry shows that

$$\frac{\sinh(\eta+\zeta)\cosh(\epsilon-\eta)}{\cosh(\epsilon-\eta-\zeta)} > \sinh\eta$$

Hence $$|P_c P_0| > |PP_0|$$

which proves the result

Since P is between P_0 and P_c, the acceleration of P, directed towards P_c, is away from P_0 This shows in a simple way why the particle P appears to be accelerated away from the fundamental particle possessing the same 'initial' velocity For P is in motion relative to P_0, so that the velocity-frame in which P is at rest does not coincide with the velocity-frame in which P_0 is at rest P_0 is the centre of the system in the frame in which P_0 is at rest, just as much as O is the centre of the system to the particle-observer O, but it is P_c which is the centre of the system in the frame in which P is at rest Since P, as we have shown, is accelerated towards P_c, and since P_c is on the opposite side of P from P_0, P's acceleration is directed away from P_0

185. The case is quite other with a particle P' which, not coinciding with P_0, is at rest relative to P_0 For such a particle, its velocity $\mathbf{V}'$ is instantaneously equal to $\mathbf{V}_0$, and the apparent centre of the system in the frame in which P' is instantaneously at rest is itself P_0 Its acceleration is therefore directed towards P_0 Of course the 'initial' velocity for this particle, which we may call momentarily $\mathbf{V}_0'$, is not equal to $\mathbf{V}_0$

To investigate this point consider two particle-trajectories passing through the same point P at the same moment, one with constants defined by $\mathbf{V}_0, \mathbf{1}, X_1$, the other with a velocity instantaneously equal to $\mathbf{V}_0$ We denote by primes (′) the symbols relating to the second trajectory Thus $\mathbf{P} = \mathbf{P}'$, $\mathbf{V}_0 = \mathbf{V}'$, $\mathbf{P}_0 = \mathbf{P}_c'$ The point P_0', by the general theory, must lie on the line joining P' and P_c', on the other side of P from P_c' It is depicted in Fig 15 Then OP_0' is the direction of V_0' The hodograph of P' must be parallel to that of P, since it is parallel to the lines $P_0 PP_c$, $P_0' P' P_c'$, which coincide Also $M' = M_0$ is a point on this hodograph Hence the hodographs coincide Hence M_0' is the point in which OP_0' meets $M_0 M$, and this defines the vector $\mathbf{V}_0' = OM_0'$ for the 'initial' velocity of P' The tangent to the trajectory of P at P is parallel to OM, the tangent to the trajectory of $P'(= P)$ at P' is parallel to $OM' = OM_0$ The acceleration of P is

towards P_c, and so away from P_0, whilst the acceleration of P' is towards P'_c, and so towards P_0 This illustrates how a particle at rest relative to P_0 at epoch t has an acceleration towards P_0, whilst a

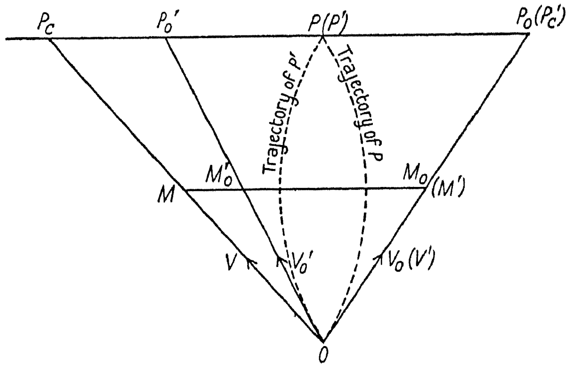


FIG 15 Pair of trajectories passing simultaneously through a point P

particle at rest relative to P_0 at the singular epoch $t = 0$ has an acceleration away from P_0

186. These relationships become perhaps clearer if we draw the diagram from the point of view of P_0 This is equivalent to taking P_0 at O The hodograph is then a straight line through O along OP The particle P, 'initially' at rest relative to O, has been accelerated away from O or P_0 to the position P, and sees the centre of the whole system as P_c, on the other side of P from O P's velocity is away from O But the particle P', in the same position as P, at rest relative to O, has zero velocity and sees the centre of the whole system as O or P_0, which is accordingly P'_c The particle P' is accelerated towards O, if we like, by the pull of the whole system centred (to P') at O, but the particle P at the same place, moving with a non-zero velocity relative to O along OP, is pulled away from O by the pull of the whole system centred (to P) at P_c, beyond P

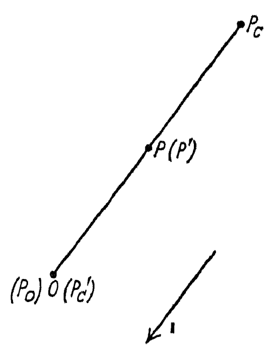


FIG 16 Trajectory seen from the associated fundamental particle

187. It is particularly to be noted that the motion or trajectory of a particle is not determined at $t = 0$ by the circumstances of projection A triple infinity of possible trajectories radiate outwards

from any point P_0 (reckoned from P_0) in straight lines, defined by arbitrary choice of X_1 and **i** If we like, we can say that at the epoch of creation there is indeterminacy The trajectory of a particle projected at any epoch $t \neq 0$ is determined by the circumstances of projection But at $t = 0$ itself these do not suffice

In Fig 15, the curved trajectory OP, viewed from the moving particle P_0, is a radial straight line radiating from P_0 In Fig 16 the curved trajectory OP is a straight line when viewed from the moving P_0, whilst the trajectory OP' is a straight line viewed from the moving P_0'

Pursuit of the apparent centre

188. We now return to Fig 14, and consider the state of affairs at a neighbouring later instant $t+dt$, reckoned by O By the time $t+dt$, P has arrived at $P+dP$, where the small arc joining P to $P+dP$ is approximately parallel to OMP_c Draw the parallel to **i** through $P+dP$, meeting OM_0P_0 in P_0+dP_0 and meeting OMP_c in S Then P_0+dP_0 is the position, at epoch $t+dt$, of the fundamental particle which at t was at P_0 At $t+dt$, let $M+dM$, on M_0M produced, be the position of the representative point in the hodograph Then the line joining O to $M+dM$ represents the velocity of the particle at $P+dP$ Consequently at $t+dt$, the apparent centre of the system in the frame in which P (at $P+dP$) is at rest is P_c+dP_c, the intersection of the line joining O to $M+dM$ with the $-$**i** transversal through P_0+dP_0

It follows that as the particle P moves from P to $P+dP$, the apparent centre of the system moves from P_c to P_c+dP_c This position is, however, occupied by a different member of the fundamental system The fundamental particle which was at P_c at epoch t has moved to S at epoch $t+dt$ Thus the apparent centre of the system has a locus amongst the fundamental particles This is because, as P is accelerated, the frame in which P is momentarily at rest alters, and so P picks out at every different epoch a different fundamental particle as occupying momentarily the centre of the system in the frame in which P is momentarily at rest

189 This has a very remarkable consequence The points P, P_c, S, $P+dP$ define a parallelogram Hence the side joining P to P_c is equal to the side joining $P+dP$ to S But $M+dM$ is beyond M, as reckoned from M_0 Hence P_c+dP_c is beyond S, as reckoned from P_0+dP_0 Hence the line joining $P+dP$ to P_c+dP_c is greater than the

line joining $P+dP$ to S, and so greater than PP_c. Hence at epoch $t+dt$ the particle is farther from the then apparent centre than it was at epoch t. The particle thus gets farther and farther away from the apparent centre of the system in the frame in which it is momentarily at rest. The particle is always 'falling freely' *towards* the apparent centre. In spite of this, so far from ever falling *into* it, its distance from it increases as time goes on. The particle P follows a curve of pursuit of P_c, but never overtakes it.

190. What we have proved, and proved rigorously, by an elementary geometrical argument, can easily be established by the sceptical reader directly from the formula for PP_c, by use of a little hyperbolic trigonometry. One of the advantages of our procedure is that it permits the valid use of simple geometrical arguments so long banished from the subject.

191. It follows that the distance of P from P_c steadily increases. It increases until M reaches M_l, when the situation requires reconsideration. When M reaches M_l, where $|OM_l| = c$, we have $|\mathbf{V}| = c$, and since always $\mathbf{P}_c$ is the vector $\mathbf{V}t$, here in the limit P_c is on the expanding frontier itself. There is no contradiction here, although it appears paradoxical to say that in the limit the apparent centre of the system is on the boundary of the system. When P approaches the velocity of light, there is still one particle which is at rest, in the limit, relative to P, although a particle actually in motion with the velocity of light has the same velocity relative to all the particles present. Although $|\mathbf{P}_c| \to ct$, P itself remains always an interior point of the system, necessarily at a finite distance inside the frontier.

192. We have still not determined whether M reaches M_l in a finite time (in O's experience) or whether M merely approaches M_l as $t \to \infty$. Clearly, the integrals for X and η, formulae (56) and (57), converge or diverge together as $\xi \to \infty$. Hence by (53), as $\xi \to \infty$, t tends to a finite limit if X and η converge to finite limits, whilst $t \to \infty$ if X and η diverge.

As M approaches M_l, whether as $t \to \infty$ or as $t \to$ a finite limit, M must come momentarily to rest. For M cannot go beyond M_l. Since the velocity of M in the hodograph is the acceleration $d\mathbf{V}/dt$ of P, it follows that as $\xi \to \infty$, $d\mathbf{V}/dt \to 0$. Now along any given trajectory

$$\frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi) = \frac{c}{X^{\frac{1}{2}}}(1-V_0^2/c^2)\frac{\sinh\zeta\cosh^2\epsilon}{\cosh^3(\epsilon-\eta-\zeta)}G(\xi)\mathbf{1},$$

and as $\xi, \zeta \to \infty$,

$$\frac{d\mathbf{V}}{dt} \sim \frac{c}{X^{\frac{1}{2}}}(1-V_0^2/c^2)\frac{\cosh^2\epsilon}{\frac{1}{4}e^{3\eta+2\zeta-3\epsilon}}G(\xi)\mathbf{i}$$

If $\eta, X \to$ finite limits as $\xi \to \infty$, then $G(\xi)e^{-2\zeta} \to 0$ as $\xi \to \infty$, i e $G(\xi)/\xi \to 0$ as $\xi \to \infty$ This gives no contradiction with the convergence of the integrals for η and X as $\xi \to \infty$, for all the conditions would be satisfied by $G(\xi) \sim$ const $\times \xi^{\alpha}$ $(0 < \alpha < 1)$ or by

$$G(\xi) \sim \text{const } \times (\log \xi)^{1+\alpha} \quad (\alpha < 0)$$

If, on the other hand, η, X diverge as $\xi \to \infty$, then

$$G(\xi)e^{-4\eta-2\zeta} \to 0 \quad \text{as} \quad \xi \to \infty,$$

i e

$$\frac{G(\xi)}{\xi}\bigg/\left[\exp -2\int^{\xi}\frac{d\xi}{\xi[1+G(\xi)]}\right] \to 0$$

as $\xi \to \infty$ As before, this sets a limit to the magnitude of $G(\xi)$ for ξ large, but tells us nothing as to the possible smallness of $G(\xi)$ Thus our present arguments carry us no further towards a decision as to whether the speed of P reaches c in a finite time or only as $t \to \infty$

Analytically it is easy to prove that P remains a finite distance inside the expanding boundary. For $P_c P = \mathbf{P}-\mathbf{V}t$ whose modulus, as we have seen, steadily increases as t increases, and as $\zeta, \xi \to \infty$, by (54)

$$\mathbf{P}-\mathbf{V}t \sim cX^{\frac{1}{2}}e^{-\eta+\epsilon}\mathbf{i},$$

whilst

$$X^{\frac{1}{2}}e^{-\eta} = \text{const } \times \exp\left[-\frac{1}{2}\int^{\xi}\left[\frac{1}{\xi-1}-\frac{1}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}}\right]\frac{d\xi}{1+G(\xi)}\right]$$

$$= \text{const } \times \exp\left[-\frac{1}{2}\int^{\xi}\frac{1-(1-\xi^{-1})^{\frac{1}{2}}}{(\xi-1)}\frac{d\xi}{1+G(\xi)}\right],$$

and this cannot approach zero since $\xi > 1$ and $1+G(\xi)$ is always negative for $\xi < \infty$

Summary of trajectory properties

193. If a free particle is projected inside the given simple kinematic system of moving particles, and collisions are ignored, its trajectory lies in a plane through the origin (any arbitrary particle-observer of the given system) containing the velocity of projection Its hodograph is a straight line The trajectory will appear in general curved, but it appears as a radial straight line judged from that fundamental particle which has moved, since $t = 0$, with the same constant velocity as the projected particle would have possessed at $t = 0$ if its trajectory

were extrapolated backwards, or continued backwards to $t = 0$ Its acceleration is radially outwards from this fundamental particle $\mathbf{V} = \mathbf{V}_0$ Its acceleration is always directed towards the apparent centre P_c of the fundamental system reckoned in the frame in which the projected particle is instantaneously at rest This point P_c is the particle which would actually be picked out by an observer moving with the projected particle as the apparent centre of spherical symmetry of the whole system at that moment The point P_c has a locus amongst the fundamental particles—it moves from fundamental particle to fundamental particle, and steadily recedes from the projected particle Thus though the projected particle P is accelerated towards P_c, and though its acceleration contains a factor proportional to $|P_c P|$, it gets steadily farther and farther away from P_c P always remains an interior particle of the system in the strict sense that its distance from the boundary does not tend to zero The process of being accelerated therefore goes on until P reaches the velocity of light, but whether it does this in a finite time or only as $t \to \infty$ we have not yet been able to determine As P approaches the speed of light, P_c tends to the boundary of the system

Incidentally we have proved several other fundamental results We have proved that there is no interaction whatever between material inside the expanding light-sphere $r = ct$ and any material which might be supposed to exist outside it Such material is therefore not only physically unobservable (owing to the impenetrable veil of material everywhere dense near the expanding frontier) but kinematically and gravitationally, and thus dynamically, unobservable Hence it is immaterial whether it be supposed to exist or not, and so irrelevant to science if a positivist point of view be adopted, propositions about it can be constructed (trajectories described, etc) but are fundamentally unverifiable

We have also proved that if we call by the name 'gravitation' the appearance of accelerations amongst particles at rest relative to a given fundamental observer, or moving with small velocities, then gravitation is necessarily a phenomenon of attraction, the accelerations are directed *towards* the local centre

Astronomical interpretation

194 If the extra-galactic nebulae are idealized to particles, and considered as on the average *equivalent* particles, then our analysis

should describe the trajectories of free particles in inter-galactic space Such a free particle, as far as we have traced its motion, continually undergoes acceleration We have obtained the accelerations from purely kinematic considerations, without adopting any 'theory of gravitation', but the accelerations undergone are compatible with the statement that every free particle recognizes an apparent centre in the whole system and is accelerated towards this centre The process of being accelerated cannot come to an end short of the particle acquiring the velocity of light, provided the particle undergoes no collisions We shall be able to prove later that actually the particle must acquire the velocity of light in a finite time, and its subsequent motion will engage our attention in great detail

This gives immediate insight into the *outward* motions suggested by the forms of spiral nebulae We have not discussed possibilities of rotation, which appear to be aspects of the imperfect way in which the nebulae, considered themselves as particles, realize 'equivalence' Our analysis must rather be regarded as dealing with the average motion of a free particle in the neighbourhood of a nebular nucleus when a large number of such motions are averaged. For example, if a large number of spiral nebulae are imagined superimposed, our analysis should give insight into the average trajectory Actually in the vicinity of a nucleus we must take account not only of a single particle but of a large number of free particles, and the development of this situation requires the statistical analysis we give in a later chapter We may be content at this stage with recognizing that for a free particle the 'centre of attraction' which it selects (to use a useful but really illegitimate mode of expression) is not necessarily the particle representing the nucleus of the neighbouring nebula, but depends on the velocity of the free particle concerned In general the free particle experiences an acceleration directed elsewhere This is not properly describable as an 'effect of velocity on acceleration', but is a purely kinematic effect associated with the structure of the complete system of fundamental particles or nebular nuclei, namely that feature of it which is the possession of an apparent centre of spherical symmetry at each fundamental particle-observer

In the space between the galaxies, complications due to congestion and distribution near a nebular nucleus do not arise, and our analysis

should adequately describe the main outlines of the motions of free particles. To the extent to which inter-galactic space is populated by matter, the particles composing such matter must be in motion and following trajectories of the kind described. The motion will be interfered with by collisions of various kinds, which will destroy momentum (to adopt here a dynamical way of speaking). The larger aggregates formed by collision will therefore be expected to possess smaller velocities than those we have calculated, though they too will undergo precisely the same accelerations, and will wander through inter-galactic space. It will be some of the smaller particles which by good fortune accidentally avoid collisions, and it is amongst these that we shall expect to find larger arcs of unimpeded trajectories, and so large velocities, even up to the velocity of light. All regions inside the expanding frontier being fully equivalent, we shall expect to find such particles equally everywhere in accessible space, independent of locality, in particular in our own vicinity.

Such particles will not have been *born* with large velocities, at creation '$t = 0$' or under any other 'initial' conditions. They will have acquired their high velocities by free fall under what may be called the gravitational field of the universe. That such particles, moving with speeds indistinguishable from that of light, are present in our own vicinity, arriving equally from all directions, is a conclusion from the observations of cosmic rays. The further discussion of the number of such particles and their relation to world-structure will be taken up later.

The principle of least action

195. As soon as the acceleration formula

$$\frac{d\mathbf{V}}{dt} = \frac{Y}{X}(\mathbf{P}-\mathbf{V}t)G(\xi) \tag{4}$$

had been obtained, the problem suggested itself of expressing this or any equivalent equation of motion under the form of the principle of least action. This problem was that of finding a function of invariants such that its integral between any two events on a trajectory was stationary for the actual path pursued. More precisely, since the only differential invariant ds is given by

$$ds^2 = dt^2 - \frac{1}{c^2}(dx^2+dy^2+dz^2),$$

the problem was to find a function of the two independent invariants X and ξ, $W(X,\xi)$, such that the equations given by

$$\delta \int W(X,\xi)\, ds = 0 \tag{60}$$

are the differential equations of the path, (4)

Since the equation (4), by its mode of construction, is of the same form for all fundamental observers, it was fairly clear that a 'least action' formulation must be possible For any equation obtained by minimalizing the integral occurring in (60) must be invariant for the same transformations The principle that the motion of a particle in *any* geometry must be capable of expression in the form of the principle of least action with suitable choice of 'weighting factor' W, a function of invariants, was shortly afterwards stated by Milner † At my request Dr A G Walker examined the problem of finding $W(X,\xi)$ so as to yield equation (4) from a variation principle of the type (60)

196. In a recent paper‡ Walker has shown that a solution to the problem exists, and he has found the most general form of the solution The solution is as follows let $f(X)$ be an arbitrary function of X, $\Theta(\sigma)$ an arbitrary function of a variable σ, let $\theta(\xi)$ be defined by

$$\theta(\xi) = \exp\left[-\int\limits_{\xi}^{\infty} \frac{d\xi}{(\xi-1)[1+G(\xi)]}\right] \tag{61}$$

Then the weighting-factor function $W(X,\xi)$ necessary to reproduce the equation of motion (4) is given by

$$W(X,\xi) \equiv X^{-\frac{1}{2}}\xi^{\frac{1}{2}} \int\limits_{f(X)}^{\xi} \Theta\{X\theta(\xi)\}(\xi-1)^{-\frac{1}{2}}\xi^{-\frac{3}{2}}\, d\xi, \tag{62}$$

X being treated as a constant in the integration with respect to ξ

If Θ is identically constant, the equation of motion (4) is reproduced only in the form

$$\frac{d\mathbf{V}}{dt} \propto \mathbf{P}-\mathbf{V}t,$$

for any other form of Θ, equation (4) is reproduced by (60) with

$$\frac{Y}{X}G(\xi) = \frac{Y}{X}\left[1+\frac{\theta(\xi)}{(\xi-1)\theta'(\xi)}\right]$$

† *Proc Roy Soc*, **139** A, 349, 1933 ‡ Ibid, **147** A, 478, 1934

It will be noticed that Walker's $\theta(\xi)$ is proportional to our integral for X, namely (9) or (56)

197. When Θ is chosen to be identically constant, $W\,ds$ is the square root of a quadratic in the differentials dx, dy, dz, dt, but the metric so obtained is degenerate, reducing to that of a space of 3 dimensions But to get the proper form of $G(\xi)$, it is shown by Walker that the case $\Theta \equiv$ const must be excluded In the general case W^2ds^2 does not reduce to a quadratic in the differentials dx, dy, dz, dt Thus though the paths of free particles may be considered as geodesics in a 4-space of metric $ds' = W\,ds$, this space is not Riemannian The space required is known as a *Finsler* space Walker concluded that 'if we wish to describe all possible natural phenomena in terms of geometry, we must use Finsler geometry, the free paths being geodesics and the light-paths null geodesics'

198. It follows that the kinematical systems here described are not equivalent to any system described by 'general' relativity, and that the kinematic systems are not included in the systems capable of description by the methods of 'general' relativity It will be shown at the end of Part IV, by totally different methods, that the simple kinematic systems do not coincide in observable properties with any of the world-systems described in current relativistic cosmology, though certain of the systems of relativistic cosmology approach the kinematic systems in the limit, as the epoch of observation (reckoned by the observer's clock) tends to infinity At any finite epoch of observation, the two sets of systems differ fundamentally, the kinematic systems possessing always an infinite number of particles in the field of view of the observer, merging into a continuous background, whilst the 'hyperbolic' systems of 'general' relativity possess always a finite (though steadily increasing) number of particles in the field of view The current cosmological solutions attain their object only at the expense of introducing a continuous creation of matter in the temporal experience of any observer, whilst the kinematic solutions contain always an infinite number of particles *already* created, i e created prior to the experience of any possible observer Since the creation of matter *in time* is contrary to our experience, we are compelled finally to reject the systems of current relativistic cosmology, whilst there is no similar objection to the kinematic solutions The important point for the moment is that two essentially different

lines of investigation concur in showing that the kinematic solutions, whilst fully consistent with the principle of relativity, are not included in that particular conceptual development of the principle which is currently known as the 'general' theory of relativity The 'general' theory of relativity is not a complete generalization of the experiences on which the principle of relativity is based This fully justifies the consideration of the new methods which are the subject of this work

IX

CONSTRUCTION OF STATISTICAL SYSTEMS

199 THE integrated form of the trajectory of a free particle, found in the preceding chapter, is fundamental in what follows We now proceed to consider cases in which a multitude of free trajectories are being described simultaneously by a large number of particles.

200. A galactic system consists of a number of particles in free motion, apart possibly from the region of the nucleus itself The free particles may be considered as stars, or as more elementary units out of which stars have been formed by collision and agglomeration We are now no longer content to idealize a nebula or galactic system as a single free particle, but wish to endow it with structure We have *a priori* no knowledge as to what structure we should posit for such a system, and we shall avoid making any specific hypothesis as to its structure Our object is in fact to ascertain on general kinematic grounds the structure of a nebula considered as an assemblage of freely moving particles

201. Our one sheet-anchor is that the different nebulae are to be considered as *equivalent* to one another An observer at the centre of any one nebula, i e situated on the free particle which idealizes its nucleus, will then describe the totality of motions constituting the totality of nebulae in the same way as another observer on any other nucleus It would, however, be altogether foreign to our line of thought to assume *a priori* that the matter of the universe is for the most part concentrated in the vicinities of the nebular nuclei We do not even know that this is the case Given the existence of matter in motion in the universe, we want to ascertain how it is distributed We have already obtained the bald outline of its distribution if it is to satisfy the cosmological principle This is described by saying that if $dxdydz$ is a volume large enough to contain a reasonably large number of particles, then the number $n\,dxdydz$ in the volume is given by

$$n\,dxdydz = \frac{Bt\,dxdydz}{c^3(t^2-\mathbf{P}^2/c^2)^2}, \tag{1}$$

where P is the vector position (x,y,z) at epoch t of any particle judged from a given particle, and that the motion of the contents of this volume is given by

$$\mathbf{V} = \mathbf{P}/t \tag{2}$$

202 It is true that we began with a set of particles constrained to move *uniformly* We then showed that they would do so naturally, without constraints Our technique, though it should be formally capable of dealing with relatively accelerated fundamental particles, has not yet been developed to the extent of embodying such motions in analysis Actually, as we shall see when we compare our results with the relatively accelerated fundamental particles described in current relativistic cosmology, accelerations of fundamental particles, if they occurred, would introduce grave physical difficulties For a particle near the expanding frontier of observable particles cannot be accelerated outwards, if its velocity is never to exceed c, if it *is* accelerated outwards, it is lost to observation, no observations on it are possible, and as far as the observable universe is concerned matter has been annihilated *in time*, in the experience of observers, which is contrary to actual experience If, on the other hand, particles are accelerated inwards, the expanding frontier can only be maintained by the creation of further particles, moving at the instant of creation with the speed of light This creation in experience is the device to which the mathematics of general relativity has recourse, at least in certain cases ($\lambda = 0$), to ensure the centrality of every particle in the field of the remainder But this amounts to creation of matter in time, in the experience of observers, which is contrary to actual experience The system of *uniformly* moving particles we have put forward avoids both annihilation and creation of matter in time, or in experience, and so is not open to the objections valid against the current relativistic theories It is therefore fully worth while to consider further a system of fundamental particles in *uniform* relative motion

203. Such particles are now to give us points of view of the further details of matter-in-motion we are going to introduce

We are no longer going to consider every particle present as a fundamental particle Instead, we are going to take a definite set of fundamental particles, defined by the value of the constant B, and surround every such particle with a distribution of other particles pursuing free trajectories and thus not necessarily unaccelerated The totality of such trajectories is now to be described in the same way as viewed from each fundamental particle This is the only method we have of avoiding preferential points of view varying from fundamental particle to fundamental particle We do not mean that

in nature every observer at the nucleus or mean centre of a nebula will see exactly the same distribution surrounding him We mean that *on the average* one point of view is as good as another, and we construct *as our model* of such motions an idealized system in which every fundamental particle does in fact see and describe the distribution in the same way

The justification of this procedure must ultimately lie in whether it gives insight into observed features of the universe and in whether it predicts the existence of other observable features It sets up a standard of comparison If its predictions are not fulfilled, then the world as a whole will not satisfy the principle postulated, and we can start again with some other model

But the model in which each fundamental particle sees the same distribution of moving particles around it is the most natural model to investigate first It is of a primary character We could only abandon the principle which denies preferential points of view by inserting preferential points of view, and to do so would be open to far more objection than to make all points of view equivalent

204. The equivalence of description is intended to hold good only for particles already equivalent to one another, in the technical sense Non-equivalent particles will not be expected to give identical points of view We are given the fundamental set of equivalent particles as a base-line It might happen that amongst the other particles surrounding them we are now to consider, pairs of equivalent particles might be present But we shall find it sufficient to fix attention on the fundamental particles

205. The set of motions defined by the fundamental particles is of hydrodynamical character At every point there is a definite velocity We have already seen (Chapter IV) that the most general hydrodynamical motion satisfying the cosmological principle is locally of the type $\mathbf{V} = \mathbf{P}F(t)$ and we are for good reasons investigating the case $\mathbf{V} = \mathbf{P}/t$ We therefore gain no further generality by considering further *hydrodynamical* motions The next simplest set of motions bears to the hydrodynamical type the same relationship as a gas bears to a liquid In a gas, in any small volume there is no unique velocity but a distribution of velocities, particles are present in the volume moving in all directions with a variety of velocities It therefore suggests itself that we should investigate the possibility of

constructing a space-velocity distribution function $f(x, y, z, t, u, v, w)$ such that

$$f(x, y, z, t, u, v, w)\, dxdydzdudvdw$$

is the number of particles present in the range $dxdydz$ near the position **P** or (x, y, z), moving with velocities inside the range $dudvdw$ near the velocity **V** or (u, v, w), all at epoch t in the experience of an assigned fundamental particle-observer

206. In attempting to discuss such a statistical set of moving particles we are essaying a far more ambitious analysis of a distribution of matter-in-motion than has previously been considered in current relativistic cosmology Existing treatments by the methods of 'general' relativity always relate to systems of hydrodynamical character, with a definite motion (relative to an assigned particle) at each definite place, inequalities or fluctuations in this motion being taken into account by introducing formally a pressure p In actual calculations, however, p seems usually to be put equal to zero, so that the systems are in the sequel of strictly hydrodynamic character But in any case no attempt has previously been made to trace a multiplicity of particle-trajectories in detail It is one of the attractions of the present method of development that it permits such an investigation with little further additional complication By the nature of the case, general relativity can only consider such detailed motions by detailed crumpling of its selected space, for the metric chosen is governed by the motions prescribed, or vice versa We continue, with equal simplicity, to retain the flat space and Newtonian time chosen by the assigned observer, and related to the spaces and times of equivalent, relatively uniformly moving observers by the formulae we have derived, namely the Lorentz formulae and their derivatives

The present method is not only more attractive but more powerful For it will enable us in due course to show that if we begin with the most general statistical distribution satisfying the cosmological principle, this is found *a posteriori* to give a set of nuclear agglomerations, and to permit a classification of the trajectories into sub-systems associated with the fundamental particles

The generalized Boltzmann equation

207. It is well known that if a distribution function f of the above type is to represent a collection of permanent objects, it must be

subject to a certain condition In the case of a gas, the condition is known as Boltzmann's equation

208. So far we have not made our distribution satisfy the cosmological principle, and for the time being we postpone doing so The particles enumerated by a given fundamental particle-observer O under the function f (obtained by simple counting) may each be undergoing some acceleration We do not so far know this acceleration But the acceleration of each particle could in principle be tabulated, and the resulting table would exhibit each acceleration written against the particle-position $\mathbf{P}$ and velocity-vector $\mathbf{V}$ identifying, at epoch t, the particle undergoing this acceleration Descriptively, then, the acceleration, say $\mathbf{g}$, is a function of $\mathbf{P}, \mathbf{V}$, and t No 'causal dependence' of acceleration on $\mathbf{P}, \mathbf{V}$, or t is implied by this function, it is purely enumerative We write then for the acceleration $\mathbf{g} = \mathbf{g}(\mathbf{P}, t, \mathbf{V})$

For brevity write f as $f(\mathbf{P}, t, \mathbf{V})$ Let do be the spatial element of volume near $\mathbf{P}$, $d\tau$ the element of velocity-range near $\mathbf{V}$, so that $f(\mathbf{P}, t, \mathbf{V})\, dod\tau$ is the number of particles counted at time t inside do with velocities inside $d\tau$

209. Consider the state of affairs at epoch $t+\Delta t$ in O's experience, Δt later than the epoch t The positions have changed to $\mathbf{P}+\Delta\mathbf{P}$, where, to a sufficient order, $\Delta\mathbf{P} = \mathbf{V}\,\Delta t$, and the velocities will have changed to $\mathbf{V}+\Delta\mathbf{V}$, where $\Delta\mathbf{V} = \mathbf{g}\,\Delta t$ Write

$$\begin{aligned} t_1 &= t+\Delta t, \\ \mathbf{P}_1 &= \mathbf{P}+\Delta\mathbf{P} = \mathbf{P}+\mathbf{V}\,\Delta t, \\ \mathbf{V}_1 &= \mathbf{V}+\Delta\mathbf{V} = \mathbf{V}+\mathbf{g}\,\Delta t, \end{aligned}$$

and let do_1 be the corresponding spatial element of volume, $d\tau_1$ the element of velocity-range Then all the particles present inside $dod\tau$ at epoch t must be present inside $do_1\, d\tau_1$ at epoch t_1 For we suppose that no particle is created or destroyed Hence

$$f(\mathbf{P}, t, \mathbf{V})\, dod\tau = f(\mathbf{P}_1, t_1, \mathbf{V}_1)\, do_1\, d\tau_1 \tag{3}$$

The algebraic details given in Note 5 show that this relation reduces to

$$\frac{\partial f}{\partial t}+\mathbf{V}\,\frac{\partial f}{\partial \mathbf{P}}+\mathbf{g}\,\frac{\partial f}{\partial \mathbf{V}}+f\left(\frac{\partial}{\partial \mathbf{V}}\,\mathbf{g}\right) = 0, \tag{4}$$

where we have used the symbol $\partial f/\partial\mathbf{P}$ to denote the vector resulting

from differentiating a scalar function f of a vector $\mathbf{P}$ partially with regard to its argument $\mathbf{P}$, similarly for $\partial f/\partial\mathbf{V}$, and $\frac{\partial}{\partial\mathbf{V}}\,\mathbf{g}$ denotes the divergence of the vector function $\mathbf{g}$ of two vectors $\mathbf{V}$ and $\mathbf{P}$ with regard to its argument $\mathbf{V}$ This is the required generalization of Boltzmann's equation which f must satisfy

Satisfaction of the cosmological principle

210. We must now arrange that the statistical system defined by the distribution function f has the same description viewed from every one of the fundamental set $\mathbf{P} = \mathbf{U}t$ This does not mean that a given particle is seen in the same place by any two fundamental observers, it means that their enumerative descriptions of the system coincide

Consider a particle at $\mathbf{P}(x, y, z)$ moving with velocity $\mathbf{V}(u, v, w)$ at epoch t as counted by O An observer O', moving with speed $(U, 0, 0)$ relative to O, reckons this event as $\mathbf{P}'(x', y', z')$, $\mathbf{V}'(u', v', w')$, t', where

$$x' = \frac{x-Ut}{(1-U^2/c^2)^{\frac{1}{2}}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t-xU/c^2}{(1-U^2/c^2)^{\frac{1}{2}}}, \quad (5)$$

$$u' = \frac{u-U}{1-uU/c^2}, \qquad v' = \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \qquad w' = \frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \quad (6)$$

and thus counts the same particle as moving at epoch t' through the position $\mathbf{P}'$ with velocity $\mathbf{V}'$

211. Consider a neighbouring particle $\mathbf{P}+d\mathbf{P}$ $(x+dx, y+dy, z+dz)$, moving with the neighbouring velocity $\mathbf{V}+d\mathbf{V}$ $(u+du, v+dv, w+dw)$ at the *same* epoch t The number of such particles inside $dod\tau$ is just $f(\mathbf{P}, t, \mathbf{V})\, dod\tau$ This particle is counted by O' at a different time $t'+\Delta t'$, in the position $\mathbf{P}'+\Delta\mathbf{P}'$ with velocity $\mathbf{V}'+\Delta\mathbf{V}'$, where

$$\Delta x' = \frac{dx}{(1-U^2/c^2)^{\frac{1}{2}}}, \qquad \Delta y' = dy, \qquad \Delta z' = dz,$$

$$\Delta t' = -\frac{(U/c^2)\,dx}{(1-U^2/c^2)^{\frac{1}{2}}}, \quad (7)$$

$$\Delta u' = \frac{du(1-U^2/c^2)}{(1-uU/c^2)^2}, \qquad \Delta v' = \frac{dv(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2} + \quad du,$$

$$\Delta w' = \frac{dw(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2} + \quad du \quad (8)$$

Hence *at the epoch t' itself*, this particle is reckoned by O' as at $\mathbf{P}'+d\mathbf{P}'$, with velocity $\mathbf{V}'+d\mathbf{V}'$, where

$$dx' = \Delta x'-u'\Delta t', \qquad dy' = \Delta y'-v'\Delta t', \qquad dz' = \Delta z'-w'\Delta t',$$
$$du' = \Delta u'-g_1'\Delta t', \qquad dv' = \Delta v'-g_2'\Delta t', \qquad dw' = \Delta w'-g_3'\Delta t',$$

where g_1', g_2', g_3' are the values, in the reckoning of O', of the components of acceleration In these formulae substitute for $\Delta x'$, , $\Delta t'$, , $\Delta w'$ in terms of dx, dy, dz, du, dv, dw from (7), (8) The results, as far as we require them, are

$$dx' = \frac{dx}{(1-U^2/c^2)^{\frac{1}{2}}}[1+u'(U/c^2)] = \frac{dx(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \tag{9}$$

$$dy' = dy+ \quad dx, \qquad dz' = dz+ \quad dx, \tag{10}$$

$$du' = \frac{du(1-U^2/c^2)}{(1-uU/c^2)^2}+ \quad dx, \qquad dv' = \frac{dv(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}+ \quad du+ \quad dx,$$

$$dw' = \frac{dw(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}+ \quad du+ \quad dx \tag{11}$$

(The coefficients denoted by are irrelevant in the sequel)

Now consider $(dx, \quad, dw)$ and $(dx', \quad, dw')$ as two sets of coordinates, the one the transform of the other Then all the particles counted by O inside $dod\tau$ at epoch t will be counted by O' as inside $do'd\tau'$ at epoch t', where

$$\frac{do'd\tau'}{dod\tau} = \frac{\partial(dx', dy', dz', du', dv', dw')}{\partial(dx, dy, dz, du, dv, dw)}$$

The only relevant terms in the Jacobian determinant occur in the principal diagonal, and are the products of the coefficients of dx, dy, dz, du, dv, dw in the first terms respectively of the above expressions (9), (10), and (11) for dx', $\,, dw'$ Hence

$$\frac{do'd\tau'}{dod\tau} = \frac{(1-U^2/c^2)^{\frac{5}{2}}}{(1-uU/c^2)^5} \tag{12}$$

The spatio-velocity particle-density is accordingly estimated by O' at time t' as

$$\frac{f\,dod\tau}{do'd\tau'} = \frac{f(\mathbf{P}, t, \mathbf{V})(1-uU/c^2)^5}{(1-U^2/c^2)^{\frac{5}{2}}}$$

If the system is to satisfy the cosmological principle, this density must be described by O' as $f(\mathbf{P}', t', \mathbf{V}')$, where f is the same function as before Hence

$$f(\mathbf{P}', t', \mathbf{V}') = f(\mathbf{P}, t, \mathbf{V})\frac{(1-uU/c^2)^5}{(1-U^2/c^2)^{\frac{5}{2}}} \tag{13}$$

This is the functional equation to be satisfied by f Two other

functional equations must be satisfied, namely those obtained by replacing $(U, 0, 0)$ by $(0, U, 0)$ and $(0, 0, U)$

To solve this functional equation, put

$$f(\mathbf{P}, t, \mathbf{V}) = Y^{-\frac{5}{2}}\phi(\mathbf{P}, t, \mathbf{V}),$$

where
$$Y = 1-\mathbf{V}^2/c^2$$

Then
$$f(\mathbf{P}', t', \mathbf{V}') = Y'^{-\frac{5}{2}}\phi(\mathbf{P}', t', \mathbf{V}'),$$

where
$$Y' = 1-\mathbf{V}'^2/c^2$$

Now
$$\frac{Y'}{Y} = \frac{1-U^2/c^2}{(1-uU/c^2)^2}$$

Hence the functional equation gives at once

$$\phi(\mathbf{P}', t', \mathbf{V}') = \phi(\mathbf{P}, t, \mathbf{V}) \tag{14}$$

and this must hold identically in $x, \quad , t, \quad , w$, together with two similar equations Hence ϕ is an invariant function of $\mathbf{P}, t, \mathbf{V}$ under Lorentz transformations of the above type, and so is of the form

$$\phi \equiv \phi\left(X, \frac{Z^2}{Y}\right) \tag{15}$$

Hence the distribution-function is of the form

$$f(\mathbf{P}, t, \mathbf{V})\, d\sigma d\tau \equiv \frac{\phi(X, Z^2/Y)}{Y^{\frac{5}{2}}}\, dxdydzdudvdw, \tag{16}$$

and this number must be independent of the scale of time and of the number chosen for c In more usual language, it is a pure number, whilst $dx \quad dw$ is of dimensions (velocity)^6(time)^3 But X is of dimensions t^2 Hence the distribution function may be written

$$\frac{\phi(X, Z^2/Y)}{c^6 X^{\frac{3}{2}} Y^{\frac{5}{2}}}\, dxdydzdudvdw, \tag{17}$$

where ϕ is of zero physical dimensions Hence $\phi(X, Z^2/Y)$ is of the form $\psi(Z^2/XY)$ or $\psi(\xi)$, and accordingly the distribution function is of the form

$$f\, dxdydzdudvdw \equiv \frac{\psi(\xi)}{c^6 X^{\frac{3}{2}} Y^{\frac{5}{2}}}\, dxdydzdudvdw \tag{18}$$

212. Insertion of this value of f in the generalized Boltzmann equation (4) gives a kinematic relation between the function ψ of a single variable ξ and the acceleration function $\mathbf{g}(\mathbf{P}, t, \mathbf{V})$

Accelerations

213 Now suppose that a statistical system of particles, with the positions and velocities distributed according to (18) for an arbitrary ψ, at an arbitrary epoch $t = t_0$, is *constrained* to move so that (18) is

always obeyed Then the acceleration $\mathfrak{g}(\mathbf{P}, t, \mathbf{V})$ of any particle is connected with ψ by equation (4), as we have seen Consider now a *free* particle in motion in the presence of this statistical system Let all the accelerations of such free particles be enumerated by each fundamental observer for all possible circumstances of projection, i e for all possible $\mathbf{P}$ and $\mathbf{V}$ for all t Then since the fundamental particles are equivalent ($O \equiv O'$), and since they describe the constrained system in equivalent ways ($O \equiv O'$), the enumerative law of accelerations as a function of $\mathbf{P}, t, \mathbf{V}$ must be the same for all O's Also the fundamental particles are in uniform relative motion Hence the result of § 101 applies, and the acceleration $\mathbf{h}$ of a free particle is of the form

$$\mathbf{h}(\mathbf{P}, t, \mathbf{V}) = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}H(\xi) \tag{19}$$

Now suppose that the constraints are removed Then every particle of the statistical system becomes a free particle, and its acceleration $\mathfrak{g}(\mathbf{P}, t, \mathbf{V})$ is of the form

$$\mathfrak{g}(\mathbf{P}, t, \mathbf{V}) = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi) \tag{20}$$

When the constraints are removed, and the given particles released, the system, however its members are accelerated, must continue to satisfy the cosmological principle For all fundamental particles are equivalent, and any change in the statistical description of the system which takes place in the experience of any one, O, must take place in the experience of any other, O', it is impossible to distinguish the experience of O from that of O'

Determinacy

214 This is of course an introduction of a principle of determinacy into the behaviour of the system, a completely irrational system might behave in any manner But the form in which we introduce such a principle of determinacy is not distinguishable from the principle of relativity itself We simply demand that the acceleration-behaviour of a system whose space-velocity distribution is described statistically in the same way by two equivalent observers is also described in the same way by the same observers This is in fact the true content of the principle of relativity That principle can either be regarded as a generalization from experiment and observation, or as a concise expression of the view that the world is a rational system as patent to human description

Aside on the principle of indeterminacy

215. I do not ignore the fact that recent developments in atomic physics have led to the formulation of a principle of indeterminacy in dynamics For such behaviour, our line of argument would break down For such behaviour, equivalent observers would not have equivalent experiences of the law governing the acceleration of other free particles, for there would be no such law—the acceleration would be indeterminate in certain phenomena It is not without significance that efforts to produce a completely relativistic scheme of quantum mechanics have so far resulted in failure Our line of thought, our embodiment of the principle of relativity in the identical experiences of equivalent observers, and our consequent derivation of the Lorentz formulae, suggests that this failure is inherent in the nature of things For in a world governed by the principle of indeterminacy, or rather *not* governed by a principle of determinacy, there are no such things as strictly equivalent observers, and the whole formulation of the theory of relativity founders That macroscopic theory which is called the theory of relativity is based essentially on a deterministic view of the behaviour of particles as exhibited in phenomena, and we cannot even begin to construct such a theory in an indeterminate or irrational world

It would be foreign to the scope of this book to pursue these ideas here, and I leave them as a suggestion to other investigators Here we are content to assume determinacy and to pursue its consequences I would, however, mention that the usual view, or the somewhat different views which find expression in Lindemann's book, rest on a sort of Kantian view of space and time as given distinct categories. 'Momentum and position', 'energy and time', cannot even be defined until ideas of time and distance have been properly formulated We have made use of no concept of distance, space, or the rigid body, but have based our formulation solely on the existence of a temporal experience for each observer How space-measures are to be constructed out of temporal experiences alone in an indeterminate world, whether indeed such a construction be possible, I have not considered. But such a consideration is imperatively demanded before a principle of indeterminacy can ever be stated It is illogical to begin by positing two categories of measures, time- and space-measures, and then to complain that they will not fit together Time

or temporal experience is the possession of each individual Space is a construct It must be reconstructed out of observations exhibiting indeterminate phenomena before such phenomena can be described in terms of space and time It is no use to say that 'space' and 'time', or our ordinary notion of them, 'break down' A time-sequence is given in experience, and analysis of experience must take this into account If the ordinary notion of 'space' breaks down, it must be replaced by something else before the concept of velocity can be mentioned or velocity measured Thus it is not surprising that no relativistic formulation of quantum mechanics has yet been successful To put the ideas of relativity into a new scheme of dynamical thought, resting on concepts derived from deterministic behaviour but fundamentally contradicting them, is to put old wine into new bottles, but to insert the quantum mechanics into relativistic theory is to attempt to put new wine into old bottles In the latter case the bottles burst, as is to be expected

Consequences of the Boltzmann equation

216 We now resume our discussion of the statistical system of particles in motion satisfying the cosmological principle Given any $\psi(\xi)$, the system thus constructed will have determinate acceleration, and so $\mathbf{g}(\mathbf{P}, t, \mathbf{V})$ and accordingly $G(\xi)$ exists The Boltzmann equation, stating a relation between $f(\mathbf{P}, t, \mathbf{V})$ and $\mathbf{g}(\mathbf{P}, t, \mathbf{V})$, gives therefore a relation between $\psi(\xi)$ and $G(\xi)$ It will be a test of the accuracy of the argumentation of § 213 that the variables $\mathbf{P}, t, \mathbf{V}$ must disappear from this relation save in the combination ξ This we now investigate

217. Inserting (18) and (20) in (4), we have, on dividing by f,

$$\left[\frac{\partial}{\partial t}+\mathbf{V}\;\frac{\partial}{\partial \mathbf{P}}+\frac{Y}{X}G(\xi)(\mathbf{P}-\mathbf{V}t)\;\frac{\partial}{\partial \mathbf{V}}\right]\log\frac{\psi(\xi)}{X^{\frac{3}{2}}Y^{\frac{3}{2}}}+\frac{\partial}{\partial \mathbf{V}}\;\left\{\frac{Y}{X}G(\xi)(\mathbf{P}-\mathbf{V}t)\right\} = 0,$$

or, say,

$$D\log\frac{\psi(\xi)}{X^{\frac{3}{2}}Y^{\frac{3}{2}}}+(\mathbf{P}-\mathbf{V}t)\;\frac{\partial}{\partial \mathbf{V}}\left\{\frac{Y}{X}G(\xi)\right\}-3t\frac{Y}{X}G(\xi) = 0,$$

where the operator D is given by

$$D \equiv \frac{\partial}{\partial t}+\mathbf{V}\;\frac{\partial}{\partial \mathbf{P}}+\frac{Y}{X}G(\xi)(\mathbf{P}-\mathbf{V}t)\;\frac{\partial}{\partial \mathbf{V}}$$

To reduce this equation, we note the following

$$\frac{\partial}{\partial t}X = 2t, \qquad \frac{\partial}{\partial \mathbf{P}}X = -\frac{2\mathbf{P}}{c^2}, \qquad \frac{\partial}{\partial \mathbf{V}}X = 0,$$

$$\frac{\partial}{\partial t}Y = 0, \qquad \frac{\partial}{\partial \mathbf{P}}Y = 0, \qquad \frac{\partial}{\partial \mathbf{V}}Y = -\frac{2\mathbf{V}}{c^2},$$

$$\frac{\partial}{\partial t}Z = 1, \qquad \frac{\partial}{\partial \mathbf{P}}Z = -\frac{\mathbf{V}}{c^2}, \qquad \frac{\partial}{\partial \mathbf{V}}Z = -\frac{\mathbf{P}}{c^2},$$

$$\frac{1}{\xi}\frac{\partial \xi}{\partial t} = \frac{2(X-tZ)}{ZX}, \qquad \frac{1}{\xi}\frac{\partial \xi}{\partial \mathbf{P}} = \frac{2(Z\mathbf{P}-X\mathbf{V})}{c^2XZ}, \qquad \frac{1}{\xi}\frac{\partial \xi}{\partial \mathbf{V}} = \frac{2(Z\mathbf{V}-Y\mathbf{P})}{c^2YZ},$$

$$DX = 2\left(t-\frac{\mathbf{V}\ \mathbf{P}}{c^2}\right) = 2Z,$$

$$DY = -\frac{2Y}{X}G(\xi)\frac{(\mathbf{P}-\mathbf{V}t)\ \mathbf{V}}{c^2} = -\frac{2Y}{X}G(\xi)(tY-Z),$$

$$D\xi = -2(\xi-1)(1+G(\xi))\frac{Z}{X},$$

$$(\mathbf{P}-\mathbf{V}t)\ \frac{\partial}{\partial \mathbf{V}}\left[\frac{Y}{X}G(\xi)\right] = -\frac{2(tY-Z)}{X}G(\xi)-\frac{2Z}{X}(\xi-1)G'(\xi)$$

Combining these we have

$$-\frac{3}{2}\ \frac{2Z}{X}+\frac{5}{2}\ \frac{2G(\xi)(tY-Z)}{X}-\frac{2\psi'(\xi)}{\psi(\xi)}\ \frac{Z}{X}(\xi-1)(1+G(\xi))-$$
$$-3t\frac{Y}{X}G(\xi)-\frac{2(tY-Z)}{X}G(\xi)-\frac{2Z}{X}(\xi-1)G'(\xi) = 0,$$

or

$$-\frac{3Z}{X}(1+G(\xi))-\frac{2\psi'(\xi)}{\psi(\xi)}(1+G(\xi))(\xi-1)\frac{Z}{X}-2(\xi-1)G'(\xi)\frac{Z}{X} = 0,$$

all terms explicit in t disappearing This gives

$$\frac{\frac{3}{2}}{\xi-1}+\frac{G'(\xi)}{1+G(\xi)}+\frac{\psi'(\xi)}{\psi(\xi)} = 0$$

We note that the only variable occurring in this relation is ξ, which confirms the correctness of the argumentation of § 213 The equation integrates at once in the form

$$(\xi-1)^{\frac{3}{2}}\psi(\xi)(1+G(\xi)) = -C,$$

where C is some constant (The negative sign is chosen for later convenience) This gives

$$G(\xi) = -1-\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)} \tag{21}$$

The law of gravitation for the statistical system

218. The relation expressed by (21) is an extremely remarkable one For it expresses what may be called 'the law of gravitation' for the system defined by $\psi(\xi)$, obtained by purely kinematic arguments

The function $\psi(\xi)$ describes fully the contents of the system It enumerates, for each volume element, how many particles are present, and how they are distributed in velocity It therefore prescribes completely the matter-in-motion present in the system Formula (21) then determines the acceleration of every particle present For it fixes $G(\xi)$ in terms of $\psi(\xi)$ and a single undetermined constant C, and so by (20) fixes the acceleration $\mathbf{g}(\mathbf{P}, t, \mathbf{V})$ In other words, formula (21) determines the accelerations to which a certain set of particles are subject, in one another's presence Nothing further can be demanded from any 'theory of gravitation' than that it prescribes the acceleration of every particle present in the system, and also the acceleration of any free particle added to the system Relation (21) achieves both For it tells us the acceleration of every particle present, and since these include all possible circumstances of projection (all $\mathbf{P}, t, \mathbf{V}$) it tells us the acceleration of any free particle supposed added to the system Actually the statistical system already contains free particles of every circumstance of projection, and nothing further is gained by considering the addition of another free particle

In the case of the simple kinematic system of fundamental particles defined by (1) and (2) we were able to determine the acceleration of every particle *already present*, this acceleration was in fact zero Thus considered as a self-contained system, this system was fully described gravitationally But we were unable fully to determine the function $G(\xi)$ describing the acceleration of an *added* free particle Now, however, we have obtained the function $G(\xi)$ in terms of the enumeration of the contents of the more general systems defined by $\psi(\xi)$

219 We have not, it is true, obtained such a formulation of gravitation that we can predict the acceleration of *any* particle forming a part of *any* system, or the acceleration of a test-particle added to such a system Thus we have not obtained a general theory of gravitation. But we have obtained the complete gravitational behaviour of a certain family of systems, namely those described by any arbitrary function $\psi(\xi)$, and that without recourse to any assumptions about the nature of gravitation, or the introduction of any empirical constants

220. We have based the derivation solely on the principle of relativity—on the possibility of realizing in nature equivalent observers in uniform relative motion. If such cannot be realized in nature, no formulation of the so-called 'special' theory of relativity is possible, and therefore no formulation of the so-called 'general' theory as a rational generalization of the special theory is possible. But there is nothing 'special' about our derivation. No restrictions have been introduced. A system, defined and described by ψ, in which particles are neither created nor destroyed, must undergo the accelerations defined by (21). Moreover, since ξ is an invariant in the experience of all the equivalent observers, the relation is invariant for transformations from the experience of any one fundamental observer to any other fundamental observer. And these are the only transformations which are relevant.

221. Our derivation of the law of gravitation for the very extensive class of systems ψ is fundamentally different from Einstein's formulation of a *general* theory of gravitation. Einstein's procedure was to assume a conceptual Riemannian space for the purposes of description of the positions and velocities of particles, defined by a metric $ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu$, and then to *posit* a linear invariant relation between the tensor $g_{\mu\nu}$, an associated tensor $G_{\mu\nu}$ derived from the tensor $g_{\mu\nu}$, and a tensor $T_{\mu\nu}$ which describes the matter present and its velocity-distribution. Such a linear relation involves two independent constants, usually called γ and λ. The 'constant of gravitation' γ merely serves to define a measure of mass, and is without significance save as to its algebraic sign. The sign of γ, and the value of λ, were chosen empirically to be consistent with the Newtonian law of gravitation. Thus, γ was chosen to be positive, and λ chosen to be zero or 'small'.

The formulation then suits the geometry, defined by the $g_{\mu\nu}$'s, to the matter present, defined by $T_{\mu\nu}$. Until the geometry has been chosen, $T_{\mu\nu}$ cannot be stated, and therein resides the conceptual element. A conceptual scheme is first formulated, in which it is posited that the matter-in-motion can be described. The rules by which it can be verified observationally that the matter is in fact present to such and such an extent, and moving in such and such a way, are formulated as a means of interpretation of the symbols used. The symbols, the conceptual element, come first, the interpretation in observation second. In addition to this conceptual element there is

the empirical element the positing of a linear relation between three tensors, and the adoption of a sign for γ and a value for λ This assumption is justified by its overwhelming success in subsuming in a single statement the observed 'laws' of dynamics and of gravitation. But save in its character of being invariant it gives no insight into *why* this statement should hold good It is not necessary, or inevitable, or the only possible restriction on possibility

222 In our treatment, on the other hand, we have abolished any conceptual element by *beginning* with observations (temporal experiences), instead of beginning with symbols and later fixing their observational interpretations We have recognized at the outset that geometry is arbitrary and made the simplification of using a Euclidian geometry for each separate observer, showing later how equivalent observers relate their temporal experiences and their constructed spaces We have introduced no empirical basis into our discussion of the accelerations undergone by particles in one another's presence, but determined them simply as the only permissible motion under the given circumstances We have been left with a single undetermined constant C, a constant thrown up by integration

Sign of the constant C

223. We now observe that the sign of C is fixed by our analysis. For the arguments that showed that, for $\xi < \infty$, $1+G(\xi)$ is negative, given in §§ 163, 173, hold good here unimpaired Hence C is positive, by (21), for $\psi(\xi)$, describing a number of particles, is essentially positive For a trajectory, if such exists, for which ξ passes through the value ∞, and then decreases again, we can say nothing so far about the sign of $1+G(\xi)$

How a statistical system may be considered as built up

224 The statistical system we are discussing can be imagined built up in the following way Consider the simple kinematic system defined by (1) and (2), with some value of B If a free test-particle be projected in the presence of this system, its acceleration is governed by some function $G(\xi)$ such that $1+G(\xi) < 0$ Now add further free particles The system immediately ceases to satisfy the cosmological principle, and the acceleration is no longer of the type (20) Continue to add further free particles, so distributed that in the end the new system once more satisfies the cosmological principle Then once

more the acceleration is given by a formula of the type (20), but with some new $G(\xi)$. The function $G(\xi)$ obtaining for the original simple kinematic system may be quite different from the $G(\xi)$ obtaining for the eventual statistical system. The two G-functions will approximate to one another the more closely the statistical system approximates to the simple kinematic system, i.e. the more closely the various particles of the statistical system congregate round or are concentrated towards the given fundamental particles forming the simple system. The extent to which this occurs will be investigated in detail shortly.

Trajectories in the statistical system

225. The general form of the trajectories of the particles in the statistical system will be identical with the form of the trajectory of a single free particle in the simple system. For they are governed by differential equations of the same form. The integrations we have carried out thus hold good as they stand, with this difference, that we now know the function $G(\xi)$, previously not determined, in terms of the distribution ψ. In the next chapter we apply these integrations to the statistical system now constructed. Since ψ fixes G (save for the magnitude of C), the whole set of trajectories pursued in the statistical system is determinate. Our object is to discuss the relationships of these trajectories.

X

THE UNIVERSE AS A GRAND SYSTEM OF NEBULAR SUB-SYSTEMS

226 We are now in a position to gain insight into the structure of the universe by comparison of its observed features with the features of the statistical system of particles in motion, constructed in the previous chapter, satisfying the cosmological principle For we now know the complete set of trajectories pursued, and the distribution of particles amongst these trajectories These trajectories can be classified according to their constants of integration, and so divided into sub-systems A particular type of division into sub-systems will be shown to correspond to the division of the material forming the universe into nebular sub-systems We shall show that the theoretical classification yields families of trajectories, the particles pursuing which are strongly concentrated towards the fundamental particles of the original simple kinematic model

Condensations or nuclear agglomerations

227. Thus far we have taken the function $\psi(\xi)$ defining the spatio-velocity distribution in the statistical system as arbitrary We shall now show that further considerations impose certain restrictions on the form of $\psi(\xi)$

We have in Chapter IX determined the form of $G(\xi)$ in terms of $\psi(\xi)$ by equation (21) of that chapter It follows that

$$-\frac{1}{(\xi-1)(1+G(\xi))} = \frac{1}{C}(\xi-1)^{\frac{1}{2}}\psi(\xi),$$

$$-\frac{1}{\xi^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}(1+G(\xi))} = \frac{1}{C}\frac{\xi-1}{\xi^{\frac{1}{2}}}\psi(\xi)$$

Hence the integrals (56) and (57), § 178, for X and η along a trajectory become

$$X = X_1 \exp\left[\frac{1}{C}\int\limits_{\xi_1}^{\xi} (\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right], \tag{1}$$

$$\eta = \frac{1}{2C}\int\limits_{1}^{\xi} \frac{\xi-1}{\xi^{\frac{1}{2}}}\psi(\xi)\,d\xi \tag{2}$$

228. X and η both steadily increase with ξ By the same argument as used in Chapter VIII, § 161, X has constantly the sign of X_1, and is never negative for observable particles As $t \to 0$ and $\xi \to 1$, since $X = t^2 - \mathbf{P}^2/c^2$, $|\mathbf{P}|$ must tend to zero, and $X \to 0$ Hence the integral

$$\int^{\xi} (\xi - 1)^{\frac{1}{2}} \psi(\xi)\, d\xi \tag{3}$$

diverges to $-\infty$ as $\xi \to 1$ Further, the integral for η tends to zero as $\xi \to 1$, and so the integral

$$\int_1 \frac{\xi - 1}{\xi^{\frac{1}{2}}} \psi(\xi)\, d\xi \tag{4}$$

converges at its lower limit These conditions set certain restrictions on the possible forms that may be chosen for $\psi(\xi)$ Further, they are mutually consistent, in that the convergence of (4) does not preclude the divergence of (3), and the divergence of (3) does not preclude the convergence of (4)

229. The divergence of (3) at $\xi = 1$ clearly requires that $\psi(\xi)$ must possess a singularity (an infinity) at $\xi = 1$ The convergence of (4) sets a limit to the severity of this singularity For example, if $\psi(\xi)$ is given a singularity of the type $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\alpha}$, then (3) requires that $\frac{3}{2} - \alpha \leqslant 0$, whilst (4) requires $2 - \alpha > 0$ Thus $\frac{3}{2} \leqslant \alpha < 2$ If $\psi(\xi)$ is given a singularity of the type $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\frac{3}{2}} \left[\log \frac{1}{\xi - 1}\right]^{\beta}$, then $\beta > -1$

It must be noticed that though $\psi(\xi)$ possesses a singularity at $\xi = 1$, the function $G(\xi)$ thus defined does not coincide with the function $G(\xi)$ appropriate to the simple hydrodynamical system alone (without the addition of the statistical set of particles) The function $G(\xi)$ defines, in the latter case, the acceleration of a single free particle, it is free from singularity at $\xi = 1$, and $G(1)$ is there a simple number If $G(1)$ is to be a simple number in the statistical case also, then

$$\lim_{\xi \to 1} \left[-1 - \frac{C}{(\xi - 1)^{\frac{3}{2}} \psi(\xi)} \right]$$

must be finite This is satisfied if, near $\xi = 1$, $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\alpha}$ ($\alpha > \frac{3}{2}$), in which case $G(1) = -1$ It is also satisfied if, near $\xi = 1$, $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\frac{3}{2}} \left[\log \frac{1}{\xi - 1}\right]^{\beta}$, provided $\beta > 0$, in which case also $G(1) = -1$ To the extent to which the simple kinematic system is an

approximation to the statistical system, we have thus proved that $G(1) \sim -1$ for the simple system, as previously conjectured, § 145

230. The singularity in $\psi(\xi)$ near $\xi = 1$ implies a congestion of matter in the vicinity of each fundamental particle of the original set $\mathbf{P} = \mathbf{V}t$ For at $\mathbf{P} = \mathbf{V}t$, $\xi = 1$ This result is of fundamental importance Without any *a priori* hypothesis, save that the system is defined to satisfy the cosmological principle, we have now proved that there must be a relatively high particle-density near each of the fundamental particles These particles, of course, have been described statistically as distributed according to the law $n = Bt/c^3(t^2-r^2/c^2)^2$ We now see that the additional free particles contemplated in the more general system are concentrated towards each fundamental particle, and of course less densely distributed in the space between the fundamental particles Thus the system possesses the 'village-community' characteristic Matter, when endowed with a variety of motions and distributed so as to satisfy the cosmological principle, must be distributed in the form of condensations, not quasi-homogeneously These condensations are not to be supposed to have formed by a *process of* condensation out of an original more uniformly distributed system They are there as part of the characteristic structure of the whole system The word 'condensation' is in fact a misnomer, the word 'agglomeration' would be better

231. The application of this result to the universe is at once apparent We have been led purely deductively to the inference that the matter of the universe will not be distributed with approximate 'uniformity', whatever that might mean, but will be gathered together into agglomeration-nuclei, as the only way of realizing obedience to the cosmological principle These nuclei are precisely the fundamental particles or nebular nuclei of the earlier idealization, when every nebula was replaced by a single particle When we no longer restrict the motion to be hydrodynamic in character, we find that the moving particles are strongly concentrated towards the nebular nuclei, in accordance with observation This shows that the simple system is an approximation to the statistical system

232. The nature of the singularity in the density near each nucleus is readily found The distribution being given in general by

$$\frac{\psi(\xi)}{c^6 X^{\frac{3}{2}} Y^{\frac{5}{2}}}\, dx dy dz du dv dw,$$

the number of particles inside a given small element of volume $dxdydz$ centred at the observer $(0, 0, 0)$ is obtained by putting $x = y = z = 0$ in the above formula and integrating over all speeds from 0 to c. To separate out the singularity, we shall first integrate over all speeds between c and some lower limit V_1. At $\mathbf{P} = 0$, $\xi = 1/(1-\mathrm{V}^2/c^2)$, $X = t^2$, and always $Y = (1-\mathrm{V}^2/c^2)$. Hence the number of particles with speeds between V_1 and c inside the element $dxdydz$ at the origin is

$$\frac{dxdydz}{c^6 t^3} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{V=V_1}^{c} \frac{\psi\{1/(1-V^2/c^2)\}}{(1-V^2/c^2)^{\frac{5}{2}}} V^2\, dV \sin\theta\, d\theta d\phi$$

$$= \frac{4\pi\, dxdydz}{c^6 t^3} \int_{V_1}^{c} \psi\left(\frac{1}{1-V^2/c^2}\right) \frac{V^2\, dV}{(1-V^2/c^2)^{\frac{5}{2}}}$$

Put
$$\frac{1}{1-V^2/c^2} = \sigma, \qquad \frac{V^2}{c^2} = 1-\frac{1}{\sigma}$$

Then
$$\frac{2V\, dV}{c^2} = \frac{d\sigma}{\sigma^2}$$

and the required number of particles is

$$\frac{2\pi\, dxdydz}{c^3 t^3} \int_{(1-V_1^2/c^2)^{-1}}^{\infty} (\sigma-1)^{\frac{1}{2}} \psi(\sigma)\, d\sigma \tag{5}$$

This gives a physical intepretation to the integral occurring in the X-integral, namely (3). Since (3) diverges as $\xi \to 1$, (5) diverges as $V_1 \to 0$. The coefficient of $dxdydz$ in (5) is the particle-density at the origin, (i.e. at any fundamental particle), of all the particles with speeds lying between V_1 and c. Since this tends to infinity as $V_1 \to 0$, there is a congestion of particles near any fundamental particle arising from the presence, in this vicinity, of the particles which have very low speeds relative to this fundamental particle.

233. To determine the distribution of particles *between* the fundamental particles, a much more delicate analysis is required. This will have our attention in Chapter XI.

Intensity of impingeing particles

234. Consider an elementary area dS at the origin $\mathbf{P} = 0$. Let $\boldsymbol{\nu}$ be a unit-vector in the direction of the normal to dS. Let us calculate the number of particles impingeing on this area in a short interval dt, from the side opposite to the direction of $\boldsymbol{\nu}$. Consider first the particles

impingeing with velocity $\mathbf{V}$ All those impingeing during the interval dt have been contained within a cylinder of volume $(\mathbf{V}\ \boldsymbol{\nu})\,dtdS$ and their number is accordingly

$$dtdS\frac{(\mathbf{V}\ \boldsymbol{\nu})\psi\{1/(1-V^2/c^2)\}}{c^6t^3(1-V^2/c^2)^{\frac{5}{2}}}dudvdw$$

This is the $\boldsymbol{\nu}$-component of the vector

$$dtdS\frac{\mathbf{V}\psi\{1/(1-V^2/c^2)\}}{c^6t^3(1-V^2/c^2)^{\frac{5}{2}}}dudvdw$$

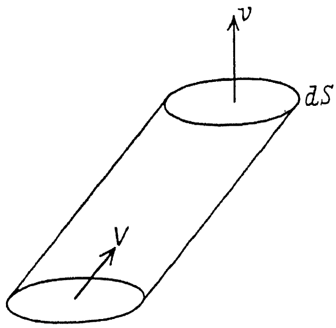


FIG 17 Calculation of intensity of the V shower

Accordingly the vector

$$\frac{1}{c^6t^3}\iiint\psi\left(\frac{1}{1-V^2/c^2}\right)\frac{\mathbf{V}}{(1-V^2/c^2)^{\frac{5}{2}}}\,dudvdw,$$

where the integral is taken over any assigned velocity range, is the 'intensity of radiation' of particles with velocities in the given range To find a typical component of this vector, take $\boldsymbol{\nu}$ along the x-axis, dS normal to the x-axis, and write $\mathbf{V} = (V\cos\theta, V\sin\theta\cos\phi, V\sin\theta\sin\phi)$ Take the range of integration to be $V_1 \leqslant V \leqslant c$ Then the number of particles in the range $V_1 \leqslant V \leqslant c$ impingeing on dS_x in the interval dt is

$$\frac{dtdS_x}{c^6t^3}\int\limits_{V_1}^{c}\int\limits_{\theta=0}^{\frac{1}{2}\pi}\int\limits_{\phi=0}^{2\pi}\psi\left(\frac{1}{1-V^2/c^2}\right)\frac{V\cos\theta}{(1-V^2/c^2)^{\frac{5}{2}}}V^2\,dV\sin\theta\,d\theta d\phi$$

$$=\frac{\pi\,dtdS_x}{c^6t^3}\int\limits_{V_1}^{c}\psi\left(\frac{1}{1-V^2/c^2}\right)\frac{V^3\,dV}{(1-V^2/c^2)^{\frac{5}{2}}}$$

Put as before

$$\frac{1}{1-V^2/c^2}=\sigma,\qquad \frac{V^2}{c^2}=1-\frac{1}{\sigma},\qquad \frac{2V\,dV}{c^2}=\frac{d\sigma}{\sigma^2}$$

Then the number is

$$\frac{\pi\, dt dS_x}{2c^2t^3} \int\limits_{(1-V_1^2/c^2)^{-1}}^{\infty} \frac{\sigma-1}{\sigma^{\frac{1}{2}}} \psi(\sigma)\, d\sigma \tag{6}$$

This integral is the integral occurring in η Since η converges at its lower limit, we may take $V_1 = 0$, and have accordingly

$$\frac{\pi\, dt dS_x}{2c^2t^3} \int\limits_{1}^{\infty} \frac{\sigma-1}{\sigma^{\frac{1}{2}}} \psi(\sigma)\, d\sigma$$

$$= \frac{\pi\, dt dS_x}{c^2t^3} C\eta(\infty), \tag{7}$$

where $\eta(\infty)$ denotes the limit of η as $\xi \to \infty$

Thus though the number of particles *present* near a fundamental particle is large, due to the excess of relatively slowly moving particles, the number impingeing on a small test-area is finite This occurs because comparatively few of the slowly moving ones have time to impinge during the given interval dt

235. But a far more important deduction can be made from (5), (6), and (7) For a system occurring in nature, the number of particles in any given vicinity moving with the speed of light must be zero The actual number with speeds V lying inside the range $V_1 \leqslant V \leqslant c$ must be finite and tend to zero as $V_1 \to c$ Hence (5) and (6) must converge at their upper limits Hence, since these same integrals re-occur in X and η, as $\xi \to \infty$, X and η must tend to finite limits This means that along any trajectory, as $\xi \to \infty$ and $V \to c$, the epoch t tends to a finite limit by formula (53), Chapter VIII Hence the point M in the hodograph (Fig 13) arrives at M_l in a finite time, and the particle P attains the velocity of light along its trajectory in a finite time By choice of the constant X_1 occurring in X and so in the formula for t, certain trajectories exist for which the value t_l (at which $|\mathbf{V}| = c$) is equal to our present epoch t It follows, since all regions are equivalent, that in every region of space, at any finite epoch, there are always some particles with speeds arbitrarily close to that of light The number of particles (or the shower-intensity) with speeds exactly equal to c is by (5) and (6) equal to zero, i e the probability of finding a particle with exactly the speed of light is zero But within any arbitrary small range, $V_1 < V < c$, close to the speed of light, there is a small non-zero number of particles

present in any volume or impingeing in a short time dt on any area We have thus obtained the answer to a question, raised in previous chapters, which we were there unable to answer The bearing of this result on cosmic rays will be considered in detail in Chapter XII

236. For the present we notice that $\psi(\xi)$ for ξ large must be such that

$$\int^{\infty} \xi^{\frac{1}{2}}\psi(\xi)\,d\xi$$

converges Hence $\xi^{\frac{3}{2}}\psi(\xi) \to 0$ as $\xi \to \infty$, and indeed $\psi(\xi) <$ const $\xi^{-\frac{3}{2}}$ as $\xi \to \infty$ This sets an upper limit to the size of $\psi(\xi)$ as $\xi \to \infty$ But a lower limit is obtained by noticing that the acceleration $\mathbf{g}$ must tend to zero as $\xi \to \infty$ For $\mathbf{g}$ may be written, since $Y = Z^2/X\xi$,

$$\mathbf{g} = (\mathbf{P}-\mathbf{V}t)\frac{Z^2}{X^2}\,\frac{1}{\xi}\left[-1-\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right],$$

whence for $\mathbf{g} \to 0$ as $\xi \to \infty$ we must have

$$\xi^{\frac{5}{2}}\psi(\xi) \to \infty$$

as $\xi \to \infty$ The two conditions are compatible They may be written

$$\frac{N}{\xi^{\frac{5}{2}}} < \psi(\xi) < \frac{\text{const}}{\xi^{\frac{3}{2}}} \qquad (\xi \sim \infty), \tag{8}$$

where N is an arbitrary large constant

Classification of trajectories

237. Our previous integration of the trajectory of a single free particle in the presence of the single kinematic system holds good for *all* trajectories in the statistical system The only difference is that η and X are now known definitely as functions of ξ in terms of the function $\psi(\xi)$ specifying the statistical distribution chosen We have found it necessary to impose certain restrictions on the behaviour of $\psi(\xi)$ for $\xi \sim 1$ and $\xi \sim \infty$, in order to obtain self-consistent statistical systems Otherwise $\psi(\xi)$ remains arbitrary Thus the results we are now about to obtain are of an extreme degree of generality We are discussing the common properties of an infinite number of possible statistical systems, the transfinite cardinal describing this number being that of the class of continuous functions of a variable Such generality has not previously been attained in any treatment of the cosmological problem In applying our results to the universe, we may say that we are discussing the common properties of an infinite number of model universes all satisfying the cosmological principle

238. It will be convenient to the reader if we repeat here the formulae for the integrated trajectories The formulae for X, η have been given as (1) and (2) above We then have (formulae (51), (52), (53), etc , § 178)

$$\mathbf{V} = \mathbf{V}_0 - c(1-V_0^2/c^2)^{\frac{1}{2}} \frac{\sinh(\eta+\zeta)}{\cosh(\epsilon-\eta-\zeta)} \mathbf{i}, \tag{9}$$

$$\mathbf{P} = \mathbf{V}_0 t - cX^{\frac{1}{2}} \frac{\sinh\eta}{\cosh\epsilon} \mathbf{i}, \tag{10}$$

$$t = \frac{X^{\frac{1}{2}}}{(1-V_0^2/c^2)^{\frac{1}{2}}} \frac{\cosh(\epsilon-\eta)}{\cosh\epsilon}, \tag{11}$$

$$\mathbf{P} - \mathbf{V}t = cX^{\frac{1}{2}} \frac{\sinh\zeta}{\cosh(\epsilon-\eta-\zeta)} \mathbf{i}, \tag{12}$$

$$Y^{\frac{1}{2}} \equiv (1-V^2/c^2)^{\frac{1}{2}} = (1-V_0^2/c^2)^{\frac{1}{2}} \frac{\cosh\epsilon}{\cosh(\epsilon-\eta-\zeta)}, \tag{13}$$

where
$$\cosh\zeta = \xi^{\frac{1}{2}}, \qquad \sinh\zeta = (\xi-1)^{\frac{1}{2}}, \tag{14}$$

and
$$\sinh\epsilon = \frac{(\mathbf{V}_0\ \mathbf{i})/c}{(1-V_0^2/c^2)^{\frac{1}{2}}} \tag{15}$$

In these formulae the origin is any arbitrary member of the fundamental system

$$n\,dxdydz = \frac{Bt\,dxdydz}{c^3X^2}, \qquad \mathbf{V} = \frac{\mathbf{P}}{t} \tag{16}$$

The integrated trajectory is expressed parametrically as a function of the single variable $\xi \equiv Z^2/XY$, and in terms of 6 arbitrary constants, $\mathbf{V}_0$ (3 constants), $\mathbf{i}$ (2 constants), and X_1 There is therefore a 6-fold infinity of such trajectories

239. Consider now the sub-set of such trajectories which have a common value of $\mathbf{V}_0$ Every trajectory of such a sub-set may be considered to have originated, at $t = 0$ ($X = 0, \eta = 0, \xi = 1$), from any arbitrary particle of the simple system (16) The arbitrary particle chosen as origin will be called O At O, at $t = 0$, the velocity along the trajectory is $\mathbf{V}_0$ There is one fundamental particle P_0 of the simple system (16) which has also started from O at $t = 0$ with the velocity $\mathbf{V}_0$ Its trajectory is simply the straight line $\mathbf{P}_0 = \mathbf{V}_0 t$, pursued with the constant velocity $\mathbf{V}_0$ We shall say that all the trajectories possessing a common value of the integration constant $\mathbf{V}_0$ form a sub-system 'based on $\mathbf{P}_0$', or 'associated with $\mathbf{P}_0$', and we shall call such a set the '$\mathbf{V}_0$-sub-system'

The '$\mathbf{V}_0 = 0$' sub-system viewed from O

240. Consider first the configuration of the $\mathbf{V}_0$-sub-system as viewed from P_0. Since all P_0's are equivalent, the view of the $\mathbf{V}_0$-sub-system from P_0 is the same as the view of the '$\mathbf{V}_0 = 0$' sub-system viewed from O. Putting $\mathbf{V}_0 = 0$, $\epsilon = 0$ in the above formulae, the '$\mathbf{V}_0 = 0$' sub-system is defined by

$$\mathbf{V} = -c\tanh(\eta+\zeta)\,\mathbf{1}, \tag{9'}$$

$$\mathbf{P} = -cX^{\frac{1}{2}}\sinh\eta\,\mathbf{1}, \tag{10'}$$

$$t = X^{\frac{1}{2}}\cosh\eta, \tag{11'}$$

$$\mathbf{P}-\mathbf{V}t = cX^{\frac{1}{2}}\frac{\sinh\zeta}{\cosh(\eta+\zeta)}\mathbf{1}, \tag{12'}$$

$$Y^{\frac{1}{2}} \equiv (1-\mathbf{V}^2/c^2)^{\frac{1}{2}} = \operatorname{sech}(\eta+\zeta) \tag{13'}$$

This is the sub-system based on O.

241. Since the direction of $\mathbf{1}$ is arbitrary, there is a sense in which, by formula (9), $\mathbf{V}_0$ may be said to be the *mean velocity* of the $\mathbf{V}_0$-sub-system. The threefold infinity of particles characterized by a given value of $\mathbf{V}_0$ have a centroid which moves with the velocity $\mathbf{V}_0$. It follows that the '$\mathbf{V}_0 = 0$' sub-system can be considered as the system 'left behind' when all the '$\mathbf{V}_0 \neq 0$' sub-systems have moved off. But the vicinity of the origin O will be populated not only by the members of the '$\mathbf{V}_0 = 0$' sub-system but also by the interpenetrating members of other ($\mathbf{V}_0 \neq 0$) sub-systems. We shall show later that the congestion near a fundamental particle $\mathbf{P}_0 = \mathbf{V}_0 t$ arises from the members of the $\mathbf{V}_0$-sub-system itself, and not from the interpenetrating members from sub-systems possessing different values of $\mathbf{V}_0$. The volume of space occupied by the members of any given $\mathbf{V}_0$-sub-system will be discussed later.

The '$\mathbf{V}_0 = 0$' sub-system, based on O, consists of members which at time $t = 0$ were at rest relative to O. But they have not remained at rest. They have been pulled out from O in different directions by the 'gravitational pull' of the whole system—to use a dynamical mode of expression. As we have seen, the event $t = 0$, $P = 0$ is a singular point on every trajectory such that the velocity there, in this case $\mathbf{V}_0 = 0$, does not uniquely define the motion. There is as it were an indeterminacy at the moment of creation.

242. According to (9'), (10'), and (11'), the '$\mathbf{V}_0 = 0$' sub-system consists of particles describing rectilinear trajectories with outwardly

accelerated velocities, each particle moving in the direction of the vector $-\mathbf{i}$ The whole sub-system possesses spherical symmetry about O For $\mathbf{V}_0 = 0$, and for a given value of $\mathbf{i}$, the various particles are strung out along the given direction at different positions given by varying X_1 Thus the threefold infinity of particles splits up into a twofold infinity of trajectory-directions $\mathbf{i}$ with a single infinity of particles along each $\mathbf{i}$ corresponding to different values of X_1 The particles for given $\mathbf{i}$, and $\mathbf{V}_0 = 0$, form a procession Their actual spatial distribution will be evaluated later

The $\mathbf{V}_0$-sub-system viewed from P_0 and from O

243 Exactly the same properties are possessed by the $\mathbf{V}_0$-sub-system as viewed from P_0 For $P_0 \equiv O$ The appearance of the $\mathbf{V}_0$-sub-system from O can be obtained by a Lorentz transformation of the appearance from P_0, i e by a transformation of equations (9′) (13′) But the result of this transformation is precisely equations (9) (13) We therefore discuss the appearance of the $\mathbf{V}_0$-sub-system as viewed from O directly by means of (9) (13)

244. The vector $\mathbf{i}$ occurs in the formulae both explicitly and in ϵ, in the form $\mathbf{V}_0\,\mathbf{i}$ It follows that viewed from O, the $\mathbf{V}_0$-sub-system possesses *axial* symmetry about the direction of $\mathbf{V}_0$ The effect of the Lorentz transformation is to compress the system by a Lorentz-contraction factor in the direction of $\mathbf{V}_0$ The sub-system possesses moreover a plane of symmetry normal to $\mathbf{V}_0$ through P_0 This may be called the equatorial plane of the $\mathbf{V}_0$-sub-system

To see that the $\mathbf{V}_0$-sub-system is compressed in the direction of $\mathbf{V}_0$, consider first the positions of particles for given ξ By (9), the absolute value of $\mathbf{P}-\mathbf{V}_0 t$ is smaller the larger is ϵ, i e the larger is $\mathbf{i}\,\mathbf{V}_0$, i e the more nearly $\mathbf{i}$ coincides with $\mathbf{V}_0$ Hence, for given ξ, the particles P are dispersed from P_0 the less the more nearly $\mathbf{i}$ is along $\mathbf{V}_0$ Thus for given ξ, the sub-system, globular when seen from P_0, is compressed in the direction OP_0 Actually for given ξ, as $\mathbf{i}$ varies t also varies, by (11), since ϵ involves $\mathbf{i}\,\mathbf{V}_0$ But a short calculation shows that the surfaces of constant t (for constant $\mathbf{V}_0$ and X_1 and varying i) are also flattened in the direction OP_0

This flattening is only appreciable when V_0 is not too small compared with c For small values of V_0/c, the sub-systems are approximately spherical in their symmetry, like the '$\mathbf{V}_0 = 0$' sub-system

245. The trajectories of the various particles forming the sub-system are obtained by taking different directions of **i** through P_0, and different values of X_1. For fixed **i** and different X_1, Fig 18 shows a number of trajectories, and the positions of the particles at a common value of ξ. The corresponding values of t are of course different, by (11), so that the particles arrive at these positions P in the diagram

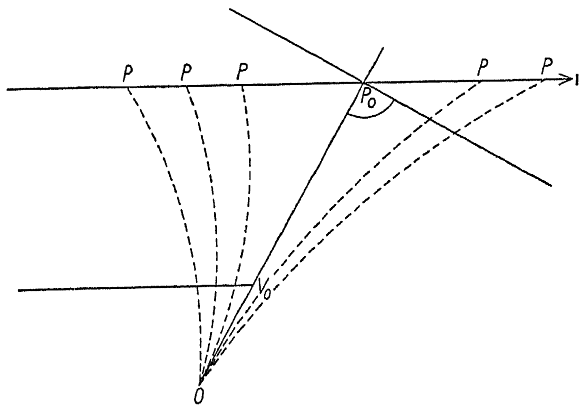


FIG 18 Trajectories with a common value of V_0

at different times t (in O's experience). Trajectories for fixed $\mathbf{V}_0$, fixed X_1, and values of **i** obtained by rotating the vector **i** about the axis of symmetry OP_0 form a flared cone, with vertex at O. As the inclination of **i** to OP_0 is varied, these flared cones expand or contract.

246. For the $\mathbf{V}_0$-sub-system based on P_0, ξ takes the value 1 for the point P_0 itself. For particles P of this sub-system near P_0, ξ is a little greater than 1, as we pass away from P_0, ξ increases. For other particles P (belonging to different $\mathbf{V}_0$-sub-systems) which happen to have moved into the vicinity of P_0, the value of ξ is not 1 or nearly 1. It follows from the singularity in $\psi(\xi)$ at $\xi = 1$, that the congestion in density near the fundamental particle $\mathbf{P}_0 = \mathbf{V}_0 t$ is due to particles of the $\mathbf{V}_0$-sub-system, and not to wanderers from other sub-systems. Hence the $\mathbf{V}_0$-sub-system is itself concentrated towards P_0, the '$\mathbf{V}_0 = 0$' sub-system is concentrated towards O.

It follows that the sixfold infinity of particle-trajectories may be decomposed into a threefold infinity of $\mathbf{V}_0$-sub-systems, each $\mathbf{V}_0$-sub-system comprising a threefold infinity of particles heavily concentrated towards the corresponding moving-particle P_0. Each sub-system is a class of outward motions, the material moving radially

outwards from the corresponding P_0 viewed from P_0 The threefold infinity of concentration-centres P_0 are distributed according to the laws (16) The various sub-systems will partially overlap one another, but the interpenetration is not serious, owing to the nuclear concentration of each sub-system In the vicinity of any nucleus there will occur wanderers from other sub-systems, but they will be in a minority Hence in the vicinity of any condensation-nucleus the predominant motion, when all the particles present are considered, will still be outwards In astronomical terminology there is a 'K-effect' in the space surrounding each nucleus—namely, an excess of radial motions, positive outwards Inter-nuclear space will be populated by particles from a variety of sub-systems, in transit away from the nuclei to which they belong

247. The nucleus to which a particle 'belongs' must not be supposed to have *generated* the particle Each particle originates equally from every nucleus, for at $t = 0$ all the nuclei were together, coinciding with an arbitrary nucleus O, and every particle has come from O, or any nucleus equivalent to O A particle of a $\mathbf{V}_0$-sub-system only 'belongs to' the corresponding P_0 in a technical sense But observation from O will at once distinguish to which $\mathbf{V}_0$-sub-systems the majority of particles 'belong', since the majority of the particles will have separated only slightly from their corresponding P_0's

Again, the nucleus to which a particle P 'belongs' must be sharply distinguished from the apparent centre P_c of the whole system in the frame of a particle-observer moving with P The position of P_c is given by $\mathbf{P}_c = \mathbf{V}t$, i e it is near that nucleus whose velocity $\mathbf{V}_0$ is most nearly equal to the instantaneous velocity $\mathbf{V}$ of P This is totally different from the $\mathbf{V}_0$ to which P 'belongs' As we saw in Chapter VIII, each P is for $\xi < \infty$ always attracted towards its P_c, and so accelerated away from its P_0 Hence the outward motions, viewed from P_0

Astronomical identification

248. We have earlier identified the nuclei of the extra-galactic nebulae with the fundamental particles $\mathbf{P}_0 = \mathbf{V}_0 t$ of the simple kinematic system (16) The properties we have now obtained for the sub-systems and their members now irresistibly compel us to identify the $\mathbf{V}_0$-sub-systems themselves with the extra-galactic nebulae considered as units possessing structure

249. The grand system of sub-systems at once reproduces the observed segregation of the material of the universe into regions of concentration near the nebular nuclei It gives the observed congestion of material near these nuclei And it reproduces the outward motions suggested by the observed forms of the spiral nebulae It does not reproduce the observed flattened forms of the spiral nebulae, or the spiral slopes of their arms These we shall discuss later The analysis must be regarded as the *average* structure of a spiral nebula that would be obtained if we superimposed a number of such nebulae, with their varying shapes and the varying orientations of their planes, and averaged them The variation of shape, form, and equatorial orientation from nebula to nebula must be regarded as a secondary effect due to the imperfect way in which a system of discretely separated nuclei satisfies the cosmological principle Our statistical, quasi-continuous analysis smooths out the local characteristics of individual nebulae, just as our hydrodynamical analysis smoothed out the general density-distribution But it gives us a great deal of information as to the *average* characteristics to be expected for a nebular sub-system

250. The grand system of sub-systems reproduces the observed expansion phenomenon with its correct sign, and the observed velocity-distance proportionality These are *nuclear* properties But it exhibits also the observed concentration of material towards these nuclei, and the outward motions from the nuclei These are structural features of individual nebular systems

Possibility of a spherical form for an individual galaxy

251. Amongst the varying forms which a sub-system may take, forms varying owing to the imperfect centrality of the system concerned in the field of the remainder, it is quite possible for some sub-system to realize approximately a perfectly central position Such a sub-system would possess spherical symmetry, and exhibit a K-effect Our own galactic system possesses an equatorial plane of symmetry But there is increasing evidence that it is more spherical in outline that it was formerly supposed to be Observations of the absorption of light in space have resulted in smaller distances being assigned to the globular clusters in low galactic latitudes, making their remoteness more nearly comparable with that of the clusters

in high galactic latitudes, and resulting in the assigning of an approximately spherical form to the system of the globular cluster as a whole † This is in accordance with our analysis

The 'K-effect'

252 Secondly, each sub-system should exhibit a 'K-effect' in radial velocities This is a well-known feature of the remoter (B-type) stars of our own system, and it here finds its explanation Members of a sub-system excentric to the nucleus are pulled away from the nucleus by what may be called the gravitational pull of the whole universe Each member is in fact accelerated towards the apparent centre of the universe in the frame in which it is momentarily at rest It would be quite possible to calculate the magnitude of this 'K-effect' From (11′), we have in fact, for the '$V_0 = 0$' sub-system, that given t means given ξ, and from (9′)

$$\frac{|\mathbf{V}|}{c} = \tanh(\eta+\zeta),$$

so that $|\mathbf{V}|$ increases with ξ (since $\eta+\zeta$ increases with ξ), and the corresponding value of $|\mathbf{P}|$ by (10′) increases with ξ Hence $|\mathbf{V}|$ increases with $|\mathbf{P}|$ But $|\mathbf{V}|$ is not proportional to $|\mathbf{P}|$ The exact form of the relation between $|\mathbf{V}|$ and $|\mathbf{P}|$, for given t, depends on $\psi(\xi)$, i e on the spatio-velocity distribution in the system For any given form of $\psi(\xi)$, at given t, (11′) determines ξ in terms of t, and then (9′) and (10′) determine $|\mathbf{V}|$ and $|\mathbf{P}|$, in terms of t

The important point for the present is that the law of expansion in a sub-system depends on the form of $\psi(\xi)$, and so is of a totally different nature from the law of expansion of the nuclei themselves, which is simply $\mathbf{P}_0 = \mathbf{V}_0 t$, independent of $\psi(\xi)$ Thus the phenomenon of the continuous expansion of a sub-system is an entirely different phenomenon from the phenomenon of the expansion of the universe Previously, opinions based on current relativistic cosmology have differed Some have held that the 'space between the galaxies expanded' alone, leaving the galaxies unaltered in dimensions, others have held that the galaxies themselves expanded with the 'space' But current relativistic cosmology has never succeeded in obtaining any precise analysis of the structure of a sub-system—it is altogether too coarse an instrument for the purpose We see that both the above

† 'The space-distribution of the globular clusters', P van de Kamp, *Astronom Journ*, **42**, no 18, May 1933

crude conjectures are wrong With us there is no question of 'expanding space', we consider simply the actual motions We then find on the average a predominantly outward motion amongst the particles forming a nebula or sub-system, of a different character from the outward motions of the nebulae or sub-systems themselves We thus offer a solution of an outstanding problem in cosmology

Whether our explanation is borne out by the observed size of the 'K-effect' in our own galaxy could only be tested by comparison of our formulae with the observed positions and mean radial velocities of the contents of our galaxy, which would serve to determine $\psi(\xi)$ Approximate formulae, assuming an appropriate form for the singularity in $\psi(\xi)$ at $\xi = 1$, could readily be worked out, as for neighbouring stars ξ can be only slightly greater than unity I have in fact made some progress with such approximations, but I shall not pause to give them here

Interlopers

253. Thirdly, we notice that on the identification here made, there must be some particles present in inter-nebular space, in transit from other nebulae Some of them will now be admixed with the particles 'belonging to' our own system Our own region of space should therefore contain (*a*) proper members of our own galaxy, possessing the value $V_0 = 0$, (*b*) foreign members which have arrived from other galaxies, $V_0 \neq 0$ To the extent to which stars represent our ideal particles, the interlopers may be called 'field stars' Although *all* trajectories originated at ourselves, or may be so considered, some of the corresponding particles have remained throughout in our own vicinity, others ('belonging to' other nebulae) have wandered away and returned Some of the latter may have much larger velocities than the velocities of 'proper' members, others, as we shall see later, may arrive with small velocities

Larmor† has recently adduced evidence that this interpenetration actually occurs in nature He has drawn attention to moving clusters which, though now inside our own system, appear to have wandered across inter-nebular space and to have entered from outside Larmor gave no explanation of the phenomenon, he simply pointed out certain facts and their dynamical implications Our analysis actually predicts‡

† *Observatory*, **57**, 55, Feb 1934

‡ I had completed my analysis and its interpretation before Larmor's welcome observational confirmation appeared

the phenomenon, predicts it moreover without recourse to a theory of gravitation, from kinematic considerations only

Thus the expansion of the universe, the velocity-distance law, the concentration of material into galaxies, the 'K-effect', the occurrence of particles in inter-nebular space and their presence by penetration within our own system are exhibited here as parts of one grand phenomenon, as essential constituents of any distribution of matter-in-motion satisfying the cosmological principle

The nature of statistical analysis

254. I feel that my interpretation of the grand statistical system, given above, may give rise to misunderstandings It may be objected that I have shut my eyes to the characteristic flattenings, spiral forms, and rotations of the nebulae and arbitrarily relegated them to secondary effects It is therefore worth while explaining in slightly more detail what our analysis is really doing

Suppose we take a 'small' volume of space, large enough to contain a number of fundamental particles, so that the formula $n\,dxdydz = Bt\,dxdydz/c^3X^2$ is validly applicable Then, in this volume of space, there are, in addition to the fundamental particles, (*a*) members of sub-systems associated with these fundamental particles, (*b*) members of sub-systems associated with fundamental particles outside the given volume Then to the particle-observer moving with the mean velocity of the fundamental particles inside the given volume, the average motion of particles of type (*a*) is given by (9′), (10′), and (11′), and these average motions define the average structure of the sub-systems associated with the fundamental particles included The average of the sub-systems therefore possesses spherical symmetry It follows that, if the individual sub-systems are supposed to be perturbing one another by local gravitational actions (not here investigated) so as to give a preferential plane for each sub-system defined by the direction of the local, residual gravitational field, then these preferential planes are oriented at random Hence in any assigned volume of space large enough to contain a reasonable number of nebular nuclei, the axes of the nebulae should be oriented at random

If now we diminish the region considered until it includes only one nucleus, our statistical analysis by its nature becomes inapplicable, for statistical methods hold good only if an elementary region of

space contains a large enough number of the entities under consideration Thus strictly speaking we can assert nothing about the form or structure of a single nebula, we can only make propositions about the mean form or structure of a large enough number It follows that we were really pushing our statistical analysis too far when we pictured all the V_0-sub-systems ($V_0 \neq 0$) as moving off and separating themselves from a single sub-system $V_0 = 0$ left behind at O All we have rigorously established is the mean form and trajectory-composition of the average of the '$V_0 \sim 0$' nebulae or sub-systems left behind in a 'small' elementary region centred at O, where the word 'small' yet implies a volume large enough to contain so many '$V_0 \sim 0$' nuclei that the statistical method remains validly applicable

The random orientation of the axes of nebulae in the small volume of space near us, predicted by our analysis, seems confirmed by observation

Collisions

255. A further point remains for discussion What do we mean by a particle? Do we mean an atom, a dust-particle, a star, or a cluster of stars?

It seems best to give no specific answer to this question, but to regard our analysis as giving the broad outlines of the distribution and trajectories of matter in motion We have throughout not taken account of the possibility of collisions These will be most frequent in a region of congestion, i e in the vicinity of the fundamental particles Hence this is where larger aggregates such as stars may be expected to be formed out of smaller units Such formations may be expected to share the mean motions of the particles out of which they are formed It is easily seen that the members of a given sub-system do not collide in the first instance amongst themselves But collisions are possible (and probable) with members of other sub-systems which have arrived in the same vicinity The aggregates so formed may then collide in turn with other members of the given sub-system, so that an aggregate may be composed of many members of the original sub-system together with some members of external sub-systems It will be observed from (7) that the frequency of impacts with a given surface of all the particles of the grand system is inversely proportional to t^3 Thus collisions are much more likely in the earlier history of each sub-system

Summary

256 The system of the nebulae may be represented as a triple infinity of sub-systems of particles, each sub-system consisting itself of a triple infinity of particles Each sub-system is characterized by a velocity $\mathbf{V}_0$, the velocity of its nucleus, $\mathbf{V}_0$ is an integration constant common to all the trajectories of the same sub-system, and prescribing the mean motion of the sub-system Each sub-system is strongly concentrated towards its nucleus which is the fundamental particle moving with uniform velocity $\mathbf{V}_0$ The members of a sub-system are in non-uniform relative motion, and undergo outward accelerations from the associated nucleus The average sub-system possesses spherical symmetry seen from its own nucleus, but the distant, relatively swiftly moving sub-systems show a Lorentz contraction in the direction of their motion The average sub-system exhibits a 'K-effect', an outward expansion, but the law expressing the rate of expansion is totally different from the law of expansion of the universe as a whole

The grand system of nebular sub-systems constitutes the universe of matter Smoothed out, it follows the simple kinematic model of our earlier analysis, and exhibits the expansion phenomenon, the velocity-distance proportionality, local homogeneity, ultimate outward increase of density, and the occupation of a finite expanding sphere of radius ct, all in the experience of any arbitrary particle-observer on an arbitrary nucleus Analysed in greater detail, it exhibits local concentration towards each fundamental, uniformly moving particle, and a local structure This local structure exhibits average spherical symmetry and local expansion

Incidentally we have proved that each particle of the grand system, save the uniformly moving fundamental particles, attains the speed of light in a finite time, thus answering an earlier question We have also obtained the total local density of non-uniformly moving particles, and their intensity of incidence on an elementary area

257. The whole of these results hold good for an arbitrary function $\psi(\xi)$ defining the details of the distribution

258 The representation of nuclear condensations which we have put forward is markedly different from that put forward in current relativistic cosmology, where the origin of condensations is attributed to

gravitational instability Hitherto it has been supposed that the 'village-community' characteristic exhibited in the aggregation of matter towards discrete points, namely the nuclei of the extra-galactic nebulae, arises from the formation of condensations in an original 'homogeneous' distribution The static homogeneous distribution known as Einstein's universe has been shown to be unstable, on certain definitions of stability and instability, and the expansion phenomenon has been held to be the consequence of the first formation of a condensation, it is clear, in fact, without the laborious calculations sometimes advanced to justify it, that should an increase of density occur locally at some point, then the matter near this point must contract, so affording the remainder of the matter the opportunity of expanding into the cavity that would otherwise be formed This view of the origin of the expansion phenomenon carries with it the consequence that each condensation should be a region of contraction Spiral nebulae show no evidence for this view from their appearance, and we here reject it, indeed, to suppose that the whole universe is expanding in all directions because a local contraction occurred at one place is to lose all sense of proportion It is obvious further that the simultaneous formation of *two* condensations in different places would cause the matter lying between them to tend to expand in two opposite directions, and there would have to be a neutral point somewhere in between where the tendencies to contract and to expand just balanced If we push the argument still further, we see that the simultaneous formation of a large number of condensations could not cause the condensation-nuclei themselves to tend to separate from one another, it would merely cause matter to drift inwards towards each condensation The average gravitational pull on each nuclear condensation due to the remainder would be unaltered Current cosmology has, in fact, got itself into great difficulties here, and authorities have sometimes differed as to whether expansion or contraction would be expected to result from the formation of condensations The methods of relativistic cosmology, with their local modifications of space-curvature for each condensation formed, are too complicated to deal satisfactorily with this topic

259. In our presentation, on the other hand, the presence of regions of congestion has been shown to be part of the very nature of a system satisfying the cosmological principle They have not been

'formed' out of an 'original' more smoothly spread distribution, but have been in existence since $t = 0$, being as it were fragments split off from the original pre-experiential singularity No question of instability arises, nor is the presence of regions of congestion directly connected with the expansion phenomenon No 'process of condensation' has 'caused' the nuclear agglomerations, the word 'condensation' suggests inward motions, but the relative motions have been invariably outward The local homogeneity in distribution of nuclei and the inhomogeneity of distribution near any individual nucleus are both consequences of the cosmological principle

In the next chapter we determine the spatial density-distribution of the particles forming a sub-system

XI

STRUCTURE OF A SUB-SYSTEM

260 WITH the warning that we are now going to push our statistical analysis to its utmost limits, and that therefore the results will not hold good for too small elementary volumes, we now proceed to investigate the density-distribution law in an average sub-system

General theorems

261. It will be convenient to preface the analysis of the structure of a sub-system with some general considerations akin to those used by Jeans and Eddington in their analysis of the steady states of star-clusters Our analysis will be, however, far more general, as it is of the essence of our whole work that we are describing non-steady states Both the grand system itself, and each separate sub-system, are examples of systems not in a steady state

Let $f(x,y,z,t,u,v,w)\,dxdydzdudvdw$ be any distribution of moving particles whose accelerations, as in our simple system and our statistical system, may be any functions of x,y,z,t,u,v,w Then the 6 differential equations of the motion giving $d\mathbf{P}/dt$ and $d\mathbf{V}/dt$ as functions of $\mathbf{P},t,\mathbf{V}$ can in principle be integrated in the form of 6 integrals

$$x = x(t, C_1, \ldots, C_6), \quad \ldots, \quad w = w(t, C_1, \ldots, C_6), \tag{1}$$

where $C_1, \ldots, C_6$ are 6 arbitrary constants of integration †

262 A domain of the C's defined by $(C_1, \ldots, C_6)$, $(C_1+dC_1, \ldots, C_6+dC_6)$ will contain a number of trajectories equal to the number of particles following them, and this number of particles will be equal to

$$f(x, \ldots, t, \ldots, w)\frac{\partial(x,y,z,u,v,w)}{\partial(C_1,C_2,C_3,C_4,C_5,C_6)_t}dC_1\,dC_2\,dC_3\,dC_4\,dC_5\,dC_6, \tag{2}$$

where the subscript t denotes that t is to be kept constant in performing the differentiations of (1) with regard to the C's We shall write this number more concisely as

$$f(\mathbf{P},t,\mathbf{V})\frac{\partial(\mathbf{P},\mathbf{V})}{\partial(C_1, \ldots, C_6)_t}dC_1 \ldots dC_6 \tag{2'}$$

But since $C_1, \ldots, C_6$ are constant along a trajectory, the number of

† It is particularly to be noted that we have nowhere had to *assume*, in our cases, that the differential equations of the motion are of the *second* order, considered as relating $\mathbf{P}$ and its time derivatives with t We actually found that accelerations could be enumerated *descriptively* in terms of $\mathbf{P}$, $\mathbf{V}$, t, and this provided the desired differential equations

trajectories with constants C_r lying inside the fixed domain $dC_1 \ldots dC_6$ cannot change as t varies. Hence

$$f(\mathbf{P}, t, \mathbf{V}) \frac{\partial(\mathbf{P}, \mathbf{V})}{\partial(C_1, \ldots, C_6)_t} \tag{3}$$

must be independent of t. It follows that if in (3) formulae (1) are substituted for $\mathbf{P}$ and $\mathbf{V}$, t must disappear from the resulting expression, which must therefore be a function of $C_1, \ldots, C_6$ only. Hence the number of particles with constants of integration lying inside $dC_1 \ldots dC_6$ is of the form

$$\phi(C_1, \ldots, C_6)\, dC_1 \ldots dC_6 \tag{4}†$$

263 Now suppose that 3 of the 6 constants, namely C_1, C_2, C_3, are chosen to lie in a prescribed domain $dC_1\, dC_2\, dC_3$. Then the distribution of particles possessing constants C_4, C_5, C_6 lying inside $dC_4\, dC_5\, dC_6$ is given by

$$dC_1\, dC_2\, dC_3\, \phi(C_1, \ldots, C_6)\, dC_4\, dC_5\, dC_6 \tag{5}$$

Now let the first three of equations (1), namely those giving $\mathbf{P}$, be solved for C_4, C_5, C_6 as functions of x, y, z, t (which of course will involve C_1, C_2, C_3) in the form

$$C_4 = C_4(x, y, z, t), \tag{6}$$

etc. Then the number of particles possessing *given* constants of integration C_1, C_2, C_3 lying inside $dC_1\, dC_2\, dC_3$ and also at time t lying inside $dxdydz$ is

$$dC_1\, dC_2\, dC_3\, \phi(C_1, \ldots, C_6) \frac{\partial(C_4, C_5, C_6)}{\partial(x, y, z)_t}\, dxdydz \tag{7}$$

In (7) we may substitute for C_4, C_5, C_6 in terms of $x, y, z, t, C_1, C_2, C_3$ from (6), and thus obtain the spatial density-distribution in the sub-system C_1, C_2, C_3 in the form

$$dC_1\, dC_2\, dC_3\, \phi_1(x, y, z, t)\, dxdydz \tag{8}$$

This may be written in the form

$$dC_1\, dC_2\, dC_3 \frac{f(\mathbf{P}, t, \mathbf{V}) \dfrac{\partial(\mathbf{P}, \mathbf{V})}{\partial(C_1, \ldots, C_6)_t}}{\dfrac{\partial(\mathbf{P})}{\partial(C_4, \ldots)_t}}\, dxdydz, \tag{9}$$

where the values of C_4, C_5, C_6 are to be substituted from (6).

† This is a far more general result than Jeans's result, *M.N., R.A.S.*, **76**, 79, 1915. He showed that in a *steady state*, $f(\mathbf{P}, \mathbf{V})$ (t absent) must be a function of the *first* integrals of the three equations of motion $d^2\mathbf{P}/dt^2 =$ function of $\mathbf{P}$. His analysis shows, in fact, that in a steady state f may be a function of the 6 integrals of this set of differential equations, but he did not make this inference. Our result (4) is not restricted to steady states. It should have many applications in the theory of star clusters.

Application to the general statistical system

264 We now apply this procedure to our statistical system of trajectories with their associated constants These have been given not as 6 functions of $t, C_1, \ldots, C_6$ but as 7 functions defining $\mathbf{P}, t, \mathbf{V}$ as functions of a parameter ξ and $\mathbf{V}_0, \mathbf{i}, X_1$, which replace $C_1, \ldots, C_6$ We can write our trajectories in the form

$$x = f_1(\xi, C_1, \ldots, C_6), \quad \ldots, \quad \ldots,$$
$$u = f_4(\xi, C_1, \ldots, C_6), \quad \ldots, \quad \ldots, \quad t = f_7(\xi, C_1, \ldots, C_6) \tag{10}$$

Then
$$\left(\frac{\partial x}{\partial C_r}\right)_t = \frac{\partial(f_1, f_7)}{\partial(C_r, \xi)} \bigg/ \frac{\partial f_7}{\partial \xi} = a_{1r},$$

say Since the distribution function with regard to $dxdydzdudvdw$ is here

$$\frac{\psi(\xi)}{c^6 X^{\frac{3}{2}} Y^{\frac{5}{2}}} dxdydzdudvdw,$$

the distribution of trajectories with respect to the C's is

$$\frac{\psi(\xi)}{c^6 X^{\frac{3}{2}} Y^{\frac{5}{2}}} ||a_{sr}|| \, dC_1 \ldots dC_6 \qquad (s, r = 1, \ldots, 6), \tag{11}$$

where $||a_{sr}||$ is the six-rowed determinant formed by the a_{sr}'s

265. By the preceding arguments, the coefficient of $dC_1 \ldots dC_6$ must be independent of t, and so of ξ It follows that $\psi(\xi)$ cannot occur in the reduced form of (11), either as it stands or in the form of an indefinite integral Thus $\psi(\xi)||a_{sr}||X^{-\frac{3}{2}}Y^{-\frac{5}{2}}$ must be a function of $C_1, \ldots, C_6$ only (independent of ξ or of $\psi(\xi)$) Hence the distribution of trajectories with respect to the C's must be given by

$$H(C_1, \ldots, C_6) \, dC_1 \ldots dC_6, \tag{12}$$

where the form of H is independent, explicitly, of the form of $\psi(\xi)$ Of course the physical meanings of the constants $C_1, \ldots, C_6$, in the sense of their relations to the events $\mathbf{P}, t, \mathbf{V}$, depend on the form of $\psi(\xi)$, but the analytical form of H will contain no mention of ψ Transformations of $C_1, \ldots, C_6$ into associated constants $C_1', \ldots, C_6'$ would alter the form of H, so that the form of H depends on the original choice of integration constants, but the number given by (12), when the elementary domain of the C's has been chosen, is independent of ψ

That (3) is of the form (4), independent of t, is a theorem of some substance That (11) is of the form (12), independent not only of ξ but also of $\psi(\xi)$, is a still more remarkable result, here of purely

kinematic character It may be considered as a further property of the integrals (9), (10), (11) of Chapter X

266. Now write $\mathbf{V}_0 = (u_0, v_0, w_0)$, $\mathbf{i} = (\cos\lambda, \sin\lambda\cos\mu, \sin\lambda\sin\mu)$† and rewrite (9), (10), (11) of Chapter X in forms giving x, y, z, u, v, w, t as functions of ξ (involving $\psi(\xi)$), $u_0, v_0, w_0, \lambda, \mu, X_1$ The value of ϵ is of course given by

$$\sinh\epsilon = \frac{(u_0\cos\lambda + v_0\sin\lambda\cos\mu + w_0\sin\lambda\sin\mu)/c}{[1-(u_0^2+v_0^2+w_0^2)/c^2]^{\frac{1}{2}}},$$

X contains X_1, which appears nowhere else, and ζ and η are functions of ξ only, η depending on $\psi(\xi)$ It now follows from the above that the trajectory distribution is

$$\frac{\psi(\xi)}{c^6 X^{\frac{3}{2}} Y^{\frac{3}{2}}} \frac{\partial(x, y, z, u, v, w)}{\partial(u_0, v_0, w_0, \lambda, \mu, X_1)_t} du_0\, dv_0\, dw_0\, d\lambda d\mu dX_1, \tag{13}$$

where the coefficient of the differential element is a function of $u_0, v_0, w_0, \lambda, \mu, X_1$ independent of ξ or of $\psi(\xi)$ after the values of x, y, z, t, u, v, w in terms of $\xi, u_0, v_0, w_0, \lambda, \mu, X_1$ have been substituted It is therefore of the form

$$H(u_0, v_0, w_0, \lambda, \mu, X_1)\, du_0\, dv_0\, dw_0\, d\lambda d\mu dX_1 \tag{14}$$

This could be verified by actual differentiation of (9), (10), (11) of Chapter X

Application to the sub-system based on the origin

267 The precise form of H in this general case I have not determined Its actual determination is only a matter of sufficient labour What we are interested in, however, is the structure of a single sub-system, and it is sufficient to detail this structure as it would be seen by an observer situated on the fundamental particle $\mathbf{P}_0 = \mathbf{V}_0 t$ on which it is based All such sub-systems will be described in the same way by the corresponding observers, since the grand system satisfies the cosmological principle Hence it is sufficient to consider the sub-system based on O, for which $\mathbf{V}_0 = 0$ For such sub-systems, $u_0 = 0$, $v_0 = 0$, $w_0 = 0$, and (14) then gives the number of particles whose associated fundamental particles lie in the small element $du_0\, dv_0\, dw_0$ centred on $u_0 = 0$, $v_0 = 0$, $w_0 = 0$

Accordingly, all the sub-systems based on particles whose velocities lie inside the small element $du_0\, dv_0\, dw_0$ centred on $u_0 = 0$, $v_0 = 0$,

† This polar angle λ is not to be confused with the λ-function used in Chapter VIII in integrating the equations of motion

$w_0 = 0$ in velocity-space contain in total a number of particles distributed according to the law

$$du_0\,dv_0\,dw_0\,H(0,0,0,\lambda,\mu,X_1)\,d\lambda d\mu dX_1 \tag{15}$$

All such sub-systems, by (9′), (10′), (11′) of Chapter X, possess spherical symmetry about O Hence the totality of particles counted in these sub-systems must be of the form

$$du_0\,dv_0\,dw_0\,H(X_1)\sin\lambda\,d\lambda d\mu dX_1, \tag{16}$$

where now $H(X_1)$ cannot depend on λ or μ The factor $\sin\lambda\,d\lambda d\mu$ fully expresses the spherical symmetry Now (16) must be a pure number, a number of particles Since X_1 is of dimensions (time)^2, $H(X_1)$ must be of the dimensions

$$\frac{1}{(\text{time})^2(\text{velocity})^3}$$

The only velocity occurring in (9′), (10′), (11′), indeed the only physical constant occurring at all, is c † Hence (16) must be of the form

$$\frac{D}{c^3}du_0\,dv_0\,dw_0\sin\lambda\,d\lambda d\mu\frac{dX_1}{X_1}, \tag{17}$$

where D is an absolute constant of zero dimensions D can therefore only be a function of the absolute constant C

268. I have actually carried out the evaluation of the function H defined by the equating of (13) and (14), by calculating the function

$$\frac{\psi(\xi)}{c^6X^{\frac{3}{2}}Y^{\frac{3}{2}}}\frac{\partial(x,y,z,u,v,w)}{\partial(u_0,v_0,w_0,\lambda,\mu,X_1)_t}$$

for the case $u_0 = v_0 = w_0 = 0$, by use of (9), (10), (11) of Chapter X The algebra is very tedious, but it was gratifying to verify that ξ and $\psi(\xi)$ completely disappeared from the resulting expression and left a formula of the type (17) The value of D comes out as

$$D = \tfrac{1}{2}C$$

This verification of a long train of argument is very satisfactory, it not only checks the accuracy of the arguments, but confirms the accuracy of integrals (9), (10), (11) of Chapter X

It follows that the total number of particles in the sub-systems based on the fundamental particles lying inside $du_0\,dv_0\,dw_0$ near

† Itself, of course, an arbitrary number fixing the relation of the length unit used to the time-unit used See Chapter II

$u_0 = v_0 = w_0 = 0$ with constants of integration λ, μ, X_1 lying inside $d\lambda d\mu dX_1$ is

$$\frac{\frac{1}{2}C}{c^3}\,du_0\,dv_0\,dw_0 \sin\lambda\,d\lambda d\mu\frac{dX_1}{X_1} \tag{18}$$

The general form of $H(u_0, v_0, w_0, \lambda, \mu, X_1)$ could now be obtained by transforming from P_0 to O. The formulae of transformation of the integration constants $\mathbf{V}_0$ and $\mathbf{i}$ for transfer of observer from P_0 to O are required. These are readily obtainable by applying a Lorentz transformation, from O to P_0, to formulae (9), (10), (11) of Chapter X, when they pass into the forms (9′), (10′), (11′) of the same chapter with a new value $\mathbf{i}'$ of '$\mathbf{i}$', defined in terms of $\mathbf{V}_0$ and the old $\mathbf{i}$. The corresponding values λ', μ' must then be inserted in (17) and the elements $du_0'\,dv_0'\,dw_0'$ and $\sin\lambda' d\lambda' d\mu'$ transformed back again from P_0 to O. I do not pause to carry out these details.

269. The number of fundamental particles lying inside $du_0\,dv_0\,dw_0$ is

$$\frac{B\,du_0\,dv_0\,dw_0}{c^3[1-(u_0^2+v_0^2+w_0^2)/c^2]^2} \tag{19}$$

Hence putting $u_0 = v_0 = w_0 = 0$, the number of fundamental particles lying inside the element $du_0\,dv_0\,dw_0$ centred on $u_0 = v_0 = w_0 = 0$ is

$$\frac{B}{c^3}\,du_0\,dv_0\,dw_0 \tag{20}$$

Consequently, dividing (18) by (20), the average number of particles, in a sub-system based on the origin, with constants of integration $\mathbf{i}, X_1$ lying inside $d\lambda d\mu dX_1$ is

$$\frac{\frac{1}{2}C}{B}\sin\lambda\,d\lambda d\mu\frac{dX_1}{X_1} \tag{21}$$

This number is independent of the choice of ψ. The meaning of X_1, it will be recalled, is that it is the value of X ($\equiv t^2-\mathbf{P}^2/c^2$) at a point on a trajectory $\mathbf{i}$ where ξ $\left(\text{that is } \left(t-\frac{\mathbf{P}\cdot\mathbf{V}}{c^2}\right)^2\left(t^2-\frac{\mathbf{P}^2}{c^2}\right)^{-1}\left(1-\frac{\mathbf{V}^2}{c^2}\right)^{-1}\right)$ takes the value ξ_1.

Formula for the spatial density-law for a sub-system

270. We can now find the spatial particle-density in a sub-system. To do so we have to pass back from the differential element $d\lambda d\mu dX_1$ to the spatial volume element $dxdydz$, and this passage reintroduces ψ. The spatial density is clearly expressed by

$$\frac{\frac{1}{2}C}{B}\frac{\sin\lambda}{X_1}\frac{\partial(\lambda, \mu, X_1)}{\partial(x, y, z)_t}\,dxdydz, \tag{22}$$

where the Jacobian is to be evaluated from the formulae (10′), (11′) of Chapter X Actual calculation gives

$$\frac{\partial(x,y,z)}{\partial(\lambda,\mu,X_1)_t} = \frac{c^3X^{\frac{3}{2}}}{2X_1}\frac{\sinh^2\eta\sinh\zeta}{\cosh(\eta+\zeta)}\sin\lambda \tag{23}$$

The Jacobian occurring in (22) is the reciprocal of (23) Inserting this, we have for the average distribution in a sub-system based on O ($\mathbf{V}_0 = 0$),

$$\frac{C}{B}\frac{\cosh(\eta+\zeta)}{\sinh^2\eta\sinh\zeta}\frac{dxdydz}{c^3X^{\frac{3}{2}}} \tag{24}$$

This, the number of particles of the sub-system inside $dxdydz$, is a pure number, as it should be

271. In (24), η and ζ are functions of ξ which are expressible in terms of x, y, z, t by means of (10′) and (11′) of Chapter X These yield

$$\tanh\eta = \frac{|\mathbf{P}|}{ct}, \tag{25}$$

so that

$$\sinh^2\eta = \frac{\mathbf{P}^2/c^2}{t^2-\mathbf{P}^2/c^2}$$

and

$$\frac{\cosh(\eta+\zeta)}{\sinh\zeta} = \frac{t}{(t^2-\mathbf{P}^2/c^2)^{\frac{1}{2}}}\left[\left(\frac{\xi}{\xi-1}\right)^{\frac{1}{2}}+\frac{|\mathbf{P}|}{ct}\right]$$

Substituting in (24) and replacing X by $t^2-\mathbf{P}^2/c^2$, we find that the spatial distribution in the sub-system $\mathbf{V}_0 = 0$ is

$$\frac{C}{B}\frac{1}{\mathbf{P}^2}\frac{1}{c^2t^2-\mathbf{P}^2}\left[ct\left(\frac{\xi}{\xi-1}\right)^{\frac{1}{2}}+|\mathbf{P}|\right]dxdydz \tag{26}$$

Substituting in (25) the definition of η as an integral involving ψ we have

$$\frac{1}{C}\int\limits_1^{\xi}\frac{\xi-1}{\xi^{\frac{3}{2}}}\psi(\xi)\,d\xi = \log\frac{ct+|\mathbf{P}|}{ct-|\mathbf{P}|} \tag{27}$$

Equation (27) is consistent with the fact that as $\xi \to 1$, $|\mathbf{P}| \to 0$, the integral on the left being known to converge If $\psi(\xi)$ is chosen or given, (27) determines ξ as a function of $|\mathbf{P}|/ct$, and then insertion of this in (26) determines the density-law (in number of particles per unit volume) in the sub-system as a function of $|\mathbf{P}|$ and t explicitly

Form of the singularity near the nucleus

272. The exact form of the density-law thus depends on the form of ψ But without further knowledge of ψ, we see that the density-distribution has a singularity at the nucleus of the sub-system,

$|\mathbf{P}| = 0$, in accordance with our previous conclusions One of the factors in this singularity is $|\mathbf{P}|^{-2}$, the other is $(\xi-1)^{-\frac{1}{2}}$ Thus the density-distribution near the nucleus varies more rapidly than the inverse square of the radial distance, $|\mathbf{P}|$ How much more rapidly depends on the form of $\psi(\xi)$ The least violent singularity permissible in $\psi(\xi)$ is $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\frac{3}{2}}$, this giving a value of $-G(1)$ not equal to 1 (More violent singularities give $-G(1) = 1$) A singularity $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\frac{3}{2}}$ inserted in (27) gives for ξ near 1, $|\mathbf{P}|/ct$ small, the relation

$$(\xi-1)^{\frac{1}{2}} \propto |\mathbf{P}|/ct,$$

which gives in (26) a singularity in density of the type $|\mathbf{P}|^{-3}$ A singularity of the type $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\frac{3}{2}-\alpha}$, $0 \leqslant \alpha < \frac{1}{2}$, which is compatible with our convergence and divergence conditions, gives $(\xi-1)^{\frac{1}{2}-\alpha} \propto |\mathbf{P}|/ct$, so that $(\xi-1)^{\frac{1}{2}} \propto [|\mathbf{P}|/ct]^{1/(1-2\alpha)}$ and the singularity in the density is of the type $|\mathbf{P}|^{-3-2\alpha/(1-2\alpha)}$, which is more severe than $|\mathbf{P}|^{-3}$ An intermediate singularity, compatible with $G(1) = -1$, is given by $\psi(\xi) \sim \text{const}\ (\xi-1)^{-\frac{3}{2}}\left(\log\frac{1}{\xi-1}\right)^{\beta}$ $(\beta > 0)$, which gives a density-singularity only logarithmically more severe than $|\mathbf{P}|^{-3}$ Thus the density-singularity is of the severity $|\mathbf{P}|^{-3}$ or higher

Projected density

273. The corresponding *projected* density, or number of particles per unit area normal to a line of sight passing a given distance l from the nucleus is readily calculated If $\rho(r)$ is the density at radial distance r, the projected density is†

$$\phi(l) = 2\int\limits_{l}^{\infty} \rho(r)\frac{r\,dr}{(r^2-l^2)^{\frac{1}{2}}},$$

which for $\rho(r) \propto r^{-3}$ gives $\phi(l) \propto l^{-2}$ near the nucleus

Comparison with observation

274 These results may be compared with the observed distribution of luminosity near the nuclei of nebulae The pioneer investigations of J H Reynolds‡ on the nucleus of the Andromeda nebula showed it to be approximately globular, as defined by isophotal lines, and obeying a luminosity law of the type *intensity* $\propto$ (distance+const)$^{-2}$.

† Plummer, *M N , R A S* , **71**, 460, 1911

‡ Reynolds, *M N , R A S* , **74**, 132, 1913, **80**, 746, 1920

Later work by Hubble† on the luminosity-distribution in elliptical nebulae confirmed Reynolds's work both in the circular character of the isophotal lines near the nucleus and the law of dependence of luminosity on radial distance

Both the spherical symmetry (implied by the circular isophotal lines) and the law of variation of luminosity with distance from the nucleus are in general agreement with the formulae found above Our formulae give, of course, an actual infinity in density at $|\mathbf{P}| = 0$ This must be supposed due to the mathematical imperfection of the neglect of collisions, for the collisions clearly cannot even approximately be ignored in the neighbourhood of the nucleus The mathematical infinity or singularity is to be taken as an indication of the possibility of relatively high densities near the nucleus In nature, the infinity must be supposed to be smoothed away to a finite density The observations both of Hubble and of Reynolds strongly suggest that this smoothing off occurs in a comparatively very restricted region near the nucleus, for their luminosity curves for the radii on opposite sides of a nucleus meet sharply at an angle without any rounding off The effect of this smoothing-out of the infinity appears to be to convert the theoretical law, of the type $\phi(l) \propto l^{-2}$ or steeper, into the less steep law $\psi(l) \propto (l+\text{const})^{-2}$, but the general agreement with our theoretical prediction seems satisfactory

Discussion

275. The determination of the density-law (26) (with ξ given by (27)) may be considered as one of the principal achievements of the kinematic theory here presented For any chosen ψ, which of course defines the populations of mixed sub-systems, (26) is perfectly definite The density-law near the nucleus for $|\mathbf{P}|/ct$ small, we have already discussed For large values of $|\mathbf{P}|$, the density falls off more slowly with distance, ultimately as $|\mathbf{P}|^{-1}(c^2t^2-\mathbf{P}^2)^{-1}$ For increasing t, at fixed $|\mathbf{P}|$, the density steadily decreases But the density is always large near the nucleus, for all t

It is remarkable that so much can be established without recourse to any dynamical or gravitational theory The kinematic method accurately takes into account what dynamically would be called both the pull of the interior material of a nuclear sub-system and the pull of the rest of the universe, external to it Of course, in the region of

† Hubble, *Astrophys Journ*, 71, 231, 1930

the nucleus itself there will be some interlopers 'belonging' to other systems, but these will be relatively sparse

276 Our analysis in this chapter is a sort of combing-out procedure We measure the density distribution in a sub-system by applying the statistical information contained in $\psi(\xi)$ to a particular sub-set of trajectories All other sub-systems, some of whose members may be present in the same volume of space considered, are as it were filtered off

We may repeat here that these regions of agglomeration have in no sense been formed by a process of condensation They are shed off, with varying velocities, from the grand singularity at $t = 0$, and owing to their differing velocities they separate from one another But each nuclear agglomeration has existed since $t = 0$ It is in process of being diluted by the outward motions of its members, but the process never ends, theoretically Each nuclear cluster of particles has the same structure, ideally, by which we mean that the average structure in any region at any time is of the type (26)

277. This system of mutually escaping sub-systems was foreseen on general grounds in my earlier paper (*Zeits fur Astrophys*, **6,** 66, 1933) There the results were anticipated by employing general ideas suggested by Newtonian gravitation Here the results have been obtained purely kinematically, without appeal to a specific theory of gravitation

Total number of particles in a sub-system

278. The total number of particles in a sub-system within given limits of ξ can be found as follows This number is

$$4\pi \int_{r'}^{r''} Nr^2\, dr,$$

where $r = |\mathbf{P}|$, and N is the coefficient of $dxdydz$ in (26) It is more convenient, however, to employ (24) The number is then

$$\frac{4\pi C}{B} \int \frac{\cosh(\eta+\zeta)}{\sinh^2\eta \sinh\zeta} \frac{r^2\, dr}{c^3 X^{\frac{3}{2}}}, \tag{28}$$

where $\qquad r = cX^{\frac{1}{2}}\sinh\eta, \qquad t = X^{\frac{1}{2}}\cosh\eta,$

and dr must be calculated for constant t, with X_1 and ξ varying appropriately We have then

$$r = ct\tanh\eta, \qquad dr = ct\,\mathrm{sech}^2\eta \frac{d\eta}{d\xi}\, d\xi,$$

whence
$$\frac{dr}{d\xi} = \frac{\psi(\xi)}{2C}\frac{cX^{\frac{1}{2}}\sinh^2\zeta}{\cosh\zeta\cosh\eta} \tag{29}$$

Then (28) reduces to
$$\frac{2\pi}{B}\int_{\xi}^{\xi''}\frac{\cosh(\eta+\zeta)\sinh\zeta}{\cosh\eta\cosh\zeta}\psi(\xi)\,d\xi$$
$$= \frac{2\pi}{B}\int_{\xi'}^{\xi'}(1+\tanh\eta\tanh\zeta)(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi \tag{30}$$

The values of ξ' and ξ'' are given in terms of $r'(=|\mathbf{P}'|)$ and $r''(=|\mathbf{P}''|)$ by means of (27)

279. By our previous results, the integral (30) diverges as $\xi'\to 1$ and converges as $\xi''\to\infty$ It follows that the total number of particles in a sub-system *including the nucleus* is infinite This is apparently a blemish on the beauty of our analysis as representing a sub-system But its origin is clear—if the total number of particles in a sub-system were finite, the sub-system would be steadily impoverished by the outward motions and would ultimately disappear Infinity of particle-number is the device to which our collisionless statistical analysis has recourse in order to ensure permanence of condensations An analysis which took into account collision processes and agglomerative processes in the congested region near a fundamental particle would simultaneously give both a finite density at the nucleus and a finite number of particles in total The singularity in (30) is of course a very mild one, the number of particles up to ξ', near the nucleus, differs by a constant from a multiple of $-\log X$, since $X = X_1\exp\Big[C^{-1}\int_{\xi_1}^{\xi}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\Big]$ Hence the infinity is logarithmic in character This can be seen otherwise by a very simple calculation for the case we have previously examined, namely $N\propto 1/r^3$ For then $\int_r Nr^2\,dr$ is of the type $\log(1/r)$ If we consider X as possessing a first-order zero near $\xi = 1$, which is all that is necessary for our analysis to be self-consistent, then this fixes the density-law (in number of particles) near a nucleus as $N\propto 1/r^3$

280. Relativistic cosmology in its current representations has not yet succeeded in dealing with the structures of a set of co-existing condensations Its method is to take a single condensation, with the

corresponding modification of geometry or metric in its neighbourhood, and smooth out the remainder Nor does current cosmology differentiate between the matter properly belonging to a sub-system and the matter properly belonging to other sub-systems present as it were incidentally in the same neighbourhood Our method has the advantage that the geometry chosen for the vicinity of a nucleus remains Euclidian The radial distance r or $|\mathbf{P}|$ from a nucleus is a length in the ordinary usage of physics, equal to that determined for example by parallax observations, by so-called rigid length-scales or by light-signals

XII

COSMIC RAYS

281. IN this chapter we proceed to consider the consequences which follow from the circumstance that the integrals for X and η must converge as $\xi \to \infty$

282. By (11), Chapter X, since X and η tend to finite limits as $\xi \to \infty$, t tends to a finite limit as $\xi \to \infty$ Since X tends to a finite limit and $X \equiv t^2 - \mathbf{P}^2/c^2$, $\mathbf{P}$ tends to a finite limit as $\xi \to \infty$ This is also evident from (10), Chapter X Since

$$\xi \equiv (t - \mathbf{P}\ \mathbf{V}/c^2)^2 (t^2 - \mathbf{P}^2/c^2)^{-1} (1 - \mathbf{V}^2/c^2)^{-1}$$

it follows that as $\xi \to \infty$, $|\mathbf{V}| \to c$ We have already commented on this astonishing fact, namely that as t increases along a trajectory, the speed of the particle reaches the speed of light in a finite time, and at a finite distance from the observer

This can be verified directly from (9), Chapter X For (9) implies (13), and as $\xi \to \infty$, $\zeta \equiv \cosh^{-1}\xi^{\frac{1}{2}} \to \infty$ and $\cosh(\epsilon - \eta - \zeta) \to \infty$ Hence $Y \to 0$, or $|\mathbf{V}| \to c$

We shall use the suffix l (the initial of *light-velocity*) to denote values of variables at the point on a trajectory at which the particle attains the speed of light Thus

$$X_l = X_1 \exp\left[\frac{1}{C}\int\limits_{\xi_1}^{\infty} (\xi - 1)^{\frac{1}{2}} \psi(\xi)\, d\xi\right], \tag{1}$$

$$\eta_l = \frac{1}{2C}\int\limits_{1}^{\infty} \frac{\xi - 1}{\xi^{\frac{1}{2}}} \psi(\xi)\, d\xi \tag{2}$$

The pure number η_l is a world-constant of the statistical system.

As a corollary,

$$X = X_l \exp\left[-\frac{1}{C}\int\limits_{\xi}^{\infty} (\xi - 1)^{\frac{1}{2}} \psi(\xi)\, d\xi\right] \tag{3}$$

and X_l may be taken as the sixth constant of integration instead of X_1 Also equation (7), Chapter X, gives for the *intensity* of particle-radiation in any direction

$$\frac{\pi\, dt dS_x}{c^2 t^3} C\eta_l \tag{4}$$

Passage through the speed of light

283 We have already noticed that the convergence of X and η to X_l and η_l implies that the representative point M in the hodograph diagram (Fig 13) reaches a point M_l distant c from O in a finite time Call the time t_l, where

$$t_l = \frac{X_l}{(1-V_0^2/c^2)^{\frac{1}{2}}} \frac{\cosh(\epsilon-\eta_l)}{\cosh\epsilon} \tag{5}$$

t_l is a function of the 6 constants $\mathbf{V}_0$, $\mathbf{l}$, and X_l Now t can certainly increase beyond t_l The fascinating question now arises, what happens to a particle for times t satisfying $t > t_l$? By choice of X_l, t_l may be made as small as we please If we identify the statistical system with the grand system of the nebular sub-systems, as we have done, then the grand system must certainly contain particles for which already, at our present epoch t, $t > t_l$ What have such particles been doing since t was equal to t_l?

284 Though we are making no use of dynamical concepts, the reader will naturally ask whether the particle has not an infinite mass when $t = t_l$ For here $|\mathbf{V}| = c$ Our answer is that the question is irrelevant For no observation on the particle can disclose the possession of an infinite mass unless the observation is made at the precise instant $t = t_l$, when $|\mathbf{V}| = c$, and the probability of an observation being made at this precise instant is zero For the number of particles in any domain possessing the exact speed $|\mathbf{V}| = c$ at any instant has been shown to be zero (§ 235, Chapter X), as $|\mathbf{V}| \to c$, the integral giving the number of particles with speeds V lying in the range $|\mathbf{V}'| < |\mathbf{V}| < c$ tends to zero, and so does the integral for the intensity of radiation of such particles Hence the probability, for example, of the particle undergoing a collision at the *exact* moment when its speed is c, is zero There is thus no contradiction with dynamics in supposing that the speed of the particle increases up to c (passing through c momentarily), if it decreases again

The mass of our whole system, in a dynamical sense, is always infinite, the system contains an infinite number of fundamental particles with speeds ranging up to just less than that of light, besides the other free particles we are now discussing The energy of the system, in the sense of dynamics, is also infinite If then it can be shown that each particle moves with the velocity of light only for a single instant $t = t_l$ of its trajectory, there will be no contradiction

with ordinary dynamics, there will be no harm in saying that an infinite amount of energy has been transferred momentarily from the total energy of the system to some particular particles of it, but the statement will have no meaning because its truth or falsehood makes no difference to any observation, it is unverifiable If indeed a collision occurred at the exact instant t_l along the trajectory at which $|\mathbf{V}| = c$, then a catastrophe might happen, but the chance of this is zero If a collision occurs *in the vicinity* of the point of the trajectory at which $|\mathbf{V}| = c$, fairly drastic effects may be expected These we consider in a moment For the present we simply notice that there is no contradiction with the ordinary principles of dynamics, including the principle of energy, provided that the particle possesses the velocity of light at a single instant t_l of its trajectory only We shall have to investigate whether this is implied by our analysis We shall find that our analysis has provided for this contingency from the start, of its own accord, and that it is not necessary to introduce any assumptions

Cosmic rays The primary agent

285. If a particle entering a neighbourhood in which there is congestion of matter possesses a velocity close to that of light, it will be likely to undergo a collision with some other material particle, and cataclysmic ionization and disintegration effects are to be expected If the neighbourhood is that of the solar system, in particular if it is that of the earth's atmosphere, the effects of the penetrating radiation that is to be expected should be observable at the earth's surface Such penetrating radiation is actually observed, it is known as cosmic radiation Let us inquire whether the primary agency producing cosmic radiation can be identified with our high-speed particles with velocities near that of light It is to be borne in mind that our high-speed particles have not been born as such, they have acquired their large velocities by freely falling, since the beginning of the world, $t = 0$, towards the apparent centre of the universe in the frame in which at each moment they are instantaneously at rest As they fall freely, they are accelerated, their pursuit of the apparent centre of the system being always 'unsuccessful' as they steadily increase in distance from the instantaneous apparent centre The sustained acceleration causes the velocity ultimately to reach that of light To speak dynamically, every such particle is being acted on by the gravi-

tational pull of the whole universe, but no theory of gravitation has been invoked in our kinematical deduction

Observed features of cosmic rays

286. The existing literature on cosmic rays, which is enormous, has been admirably summarized in A Corlin's recent monograph 'Cosmic ultra-radiation in Northern Sweden',† from which the following account is taken, and to which I desire to make full acknowledgements

Corlin remarks 'The discovery of an ultra-penetrating radiation followed as a consequence of the investigations of ionization in a closed vessel ' The researches of C T R Wilson and H Geitel in 1901 showed that ions are produced in the gas thus contained, rendering the gas slightly conducting It was at first supposed that the source of the ions was the radioactive deposits in the soil, and the ionization was attributed to the γ-radiation they emitted, for surrounding the vessel by a thick lead sheath failed to diminish the ionization perceptibly The ionization did not, however, decrease with height in the expected way Balloon flights by V P. Hess in 1911 up to over 5 km and by Kolhorster in 1913–14 up to over 9 km showed that the ionization increased upwards from about 700 m height, reaching at 9 km a value forty times higher than the ground value They concluded that a highly penetrating radiation traversed the earth's atmosphere, and ultimately proposed the name 'cosmic ultra-radiation' or simply 'cosmic radiation' Experiments on ionization in chambers sunk to great depths in the water of mountain lakes free from radioactive contamination confirmed the existence of the phenomenon, which from 1926 onwards has been generally accepted

Further experiments disclosed a 'barometer effect' Efforts were made to determine whether the culmination of the Milky Way caused any increase in ionization, but no such correlation was found No certain indication of a dependence of observed ionization effects on sidereal time has ever been established This is a very important negative result, as it suggests that the origin of the cosmic radiation cannot be associated with the direction of the centre of our own galaxy A slight dependence on mean solar time has, however, been found by many investigators, but the effect is far too small to justify an attribution of the origin of the whole of the radiation to the sun Moreover, the radiation persists during solar eclipses

† *Annals of the Observatory of Lund,* No 4, 1934

Up to 1928 it was generally supposed that the radiation was of the nature of γ-radiation, and Millikan and others associated it with the hypothetical formation of nuclei of helium, oxygen, and silicon out of protons and electrons in inter-galactic spaces, with emission of the energy corresponding to the mass-differences as single quanta of radiation. But Rutherford pointed out that it might be due to high-energy electrons entering the earth's atmosphere. The invention of the Geiger-Muller counter made it possible to seek experimental confirmation of this hypothesis. By arranging two counters in line in such a way that coincident impacts from the same entering electron could be recorded, Bothe and Kolhorster carried out observations from which they concluded that the coincidences were caused by corpuscles which had the same penetrating power as that previously assigned to the cosmic radiation, and they inferred (1929) that cosmic ultra-radiation is mainly or wholly of corpuscular character.

Corlin summarizes the history of the experiments on cosmic radiation as thus divided into three phases: a first phase, in which decision was sought as to whether the radiation was terrestrial or extra-terrestrial in origin, a second phase, in which the question at issue was the possibility of a dependence on sidereal time, and a third phase, in which the question was as to the corpuscular nature of the radiation.

Important experiments on the intensity of cosmic radiation were made by Regener, who sank instruments down to a depth of 230 m. in Lake Constance, and sent up self-registering apparatus in balloons up to 27 km. in the atmosphere. Regener inferred that the cosmic radiation carried a flux of energy equal to 3.5×10^{-3} erg. sec.$^{-1}$ cm.$^{-2}$, which is of the same order of magnitude as the energy arriving at the earth from the stars in the form of light and heat.

The observed cosmic radiation consists of 'softer' and 'harder' components, whose mass-absorption coefficients at different altitudes were measured by Kolhorster. He found the mass-absorption coefficient to have a maximum between 6 km. and 7 km., and he concluded that this maximum is caused by the increasing saturation of secondary radiation downwards and by decreasing intensity of primary radiation.

Balloon ascents to 16 km. have been made by Picard and Cosyns and the ionization measured.

The theory of Stormer and the possible identification of the radiation with swift charged corpuscles analogous to those supposed by Stormer and others to cause the auroral and magnetic storms led to

investigations to determine whether there was a dependence of intensity on geomagnetic latitude. Important measures were made by J. Clay in three voyages between Genoa and Java (the third in 1929), and he found a decrease of intensity towards the equator. Farther north (Norway and the Baltic) a decrease was again found, by Corlin. This implication of a zone of maximum intensity of cosmic radiation was investigated by coordinated researches organized by A. H. Compton and others, at various altitudes and at latitudes ranging from the far north to the far south. These observations confirmed Clay's results, and showed a minimum of cosmic ray intensity at the geomagnetic equator, the intensity increasing northward and southward. This latitude effect was more pronounced the higher the altitude, the increase of intensity at 50° N. and S. geomagnetic latitude being 14 per cent. at sea-level, 22 per cent. at 2 km. altitude, and 33 per cent. at 4.36 km. altitude, all compared with the equatorial intensity. Compton proved also that the variation was best correlated with geomagnetic latitude and not with geographic latitude or local magnetic latitude. Compton therefore concluded that the deflexion of charged corpuscles in the earth's magnetic field, predicted by Stormer's theory, is 'not only valid for the secondary corpuscles liberated in the atmosphere', but 'inherent in the primary radiation which must come from remote spaces'. The work of Compton and Clay definitely established the cosmic origin of ultra-radiation, and showed that it consists—at least to a considerable extent—of charged particles. The increase of latitude effect with altitude was ascribed to the soft component, which according to Regener constitutes more than 98 per cent. of cosmic ultra-radiation at the 'top of' the atmosphere.

Stormer's theory was applied by Clay to calculate the minimum energy of the corpuscles required to penetrate the earth's atmosphere in a given latitude. If a 'system of corpuscles of all energies from zero to infinity be incident on the Earth', then there are two factors controlling the possibility of incidence on the earth's surface, namely Stormer's 'forbidden spaces' at which particles of too low energy cannot arrive in low geomagnetic latitudes, and the earth's atmosphere, which of course must be penetrated. The fact that rays arrive at the magnetic equator was shown to imply that the arriving corpuscles must possess at least 10^{10} electron-volts of energy. This was confirmed by the three-counter experiments of Rossi.

Wilson-chamber photographs taken in a strong magnetic field ($\sim$ 20,000 gauss), which were obtained by C D Anderson, not only disclosed a small percentage of tracks with a curvature indicating energies up to 10^{10} electron-volts for the cosmic corpuscles but led to the discovery of the positive electron, or positron, a new fundamental particle with a positive charge and a mass equal or approximately equal to that of the electron This work was confirmed in great detail by Blackett and Occhialini It was concluded that cosmic radiation includes positrons

It thus became important to investigate the predominant sign of the charge of the corpuscular radiation It was shown by Lemaître and Vallarta in 1933, following Stormer's methods, that the sign of the charge determines the distribution in azimuth of the charge arriving at the earth after traversing the earth's magnetic field Primary corpuscles must arrive within a cone whose axis is directed to the west for positive particles and to the east for negative corpuscles They were, however, really rediscovering a consequence of Stormer's theory already pointed out by Rossi in 1930 The preponderance of accurate observations now show that there is a predominance in intensity towards the west, thus indicating that the cosmic radiation consists to at least a great part of *positively* charged particles, which we may identify with positrons The west-east intensity-difference increases towards the magnetic equator and increases with altitude, at 3 km height, the excess of west-arriving radiation is 13 per cent at the geomagnetic equator, 7 per cent at latitude 29°, 7 per cent at latitude 48°, in terms of the mean intensity, for an inclination of 45° to the zenith, at the equator the excess is 7 per cent at sea-level, 10 per cent at 4 2 km Complications in azimuthal distribution have been found which may be interpreted as effects of secondary radiation, but there is also evidence that a part of the primary radiation may consist of negative corpuscles

In addition to the more usual types of tracks found in Wilson-chamber photographs, extraordinarily complicated groups or bursts of tracks have been found, consisting of 'showers' of atomic nuclei It has been shown that the radiation responsible for these bursts or showers must be extremely penetrating, and that the total energy of a large single-burst (which may result in the production of a very large number of disintegration products of various kinds) may amount to 10^{11} electron-volts The shower-photographs of Blackett

and Occhialini may especially be mentioned Since the publication of Corlin's summary, Gilbert working with Blackett has shown that the effects observed in ionization chambers indicate the existence of three types of radiation, a primary radiation which is *probably* charged, a shower-producing radiation, and lastly the shower-particles themselves

Corlin summarizes the facts at present known by saying that we have to do with a penetrating radiation, primarily of cosmic origin, comprising at least six different components harder than the hardest known γ-rays from radioactive substances, the hardest is capable of penetrating 600 metres of water, and has a maximum energy exceeding 10^{11} electron-volts A considerable part of this primary radiation consists of positrons, and the whole of the radiation may be corpuscular This passage through the earth's magnetic field and consequent deflexion of direction gives rise to the 'latitude effect' (absence of particles of higher energy from regions of low geomagnetic latitude) and the 'azimuthal effect' (excess of west-arriving particles due to the excess of positive corpuscles) The intensity shows also a 'barometer effect' and a diurnal period, but no certain dependence on sidereal time

Theories of the origin of cosmic radiation

287. Various theories have been proposed for the origin of ultra-radiation C T R Wilson has attributed it to the effects of thunderstorms and the resulting acceleration of runaway electrons under the influence of the enormous potential differences brought into being. It is quite possible that the phenomena observed include the effects of such accelerated electrons, but the general opinion seems to be that the major part of the ultra-radiation is of cosmic origin

Other groups of theories have been based on the possibility of the generation of penetrating radiation in the interiors of stars, or by processes of nuclear annihilation or synthesis in interstellar space But however penetrating the radiation, it can hardly come from the far interiors of stars, and on the other hand conditions in the outer layers are not sufficiently drastic for the generation of radiation of such energy, for the temperatures (up to a few hundred thousand degrees, say, in the available layers) correspond to a few electron-volts The Nova phenomenon (probably the collapse of a star from an 'ordinary' diffuse state to a very dense state) might be capable of generating

radiation of sufficient energy, but this is unlikely † Further, the occurrence of novae in our own galaxy is scarcely sufficiently frequent, and moreover, if it were, there should be an excess of penetrating radiation from the direction of the centre of our own galaxy, which is not observed On the whole, novae could only generate cosmic rays of the observed energy if the nucleus of the star were exposed to free space during the outburst, whilst the evidence is that the outburst or collapse merely results in the throwing off, under radiation pressure, of the outer, atmospheric layers of the collapsing star

The earlier hypothesis of Millikan, attributing cosmic radiation to synthesis of heavier atoms by the simultaneous collision of sufficient protons and electrons with emission of the energy equivalent to the mass-defect in the form of photons, encounters the difficulty that such collision phenomena would be expected to be excessively rare, besides the fundamental difficulty that the radiation is now held to be corpuscular The latter objection also rules out the annihilation hypothesis of Jeans

288 Corlin gives it as his opinion that 'the present situation is quite desperate with respect to the possibility of forming any well-founded idea about the probable origin of cosmic ultra-radiation', but, as he remarks, the origin of cosmic ultra-radiation must exist somewhere He alludes, however, to a possible hypothesis mentioned by Bothe and Kolhorster,‡ without discussing it seriously This is to the effect that 'cosmic ultra-radiation consists of corpuscles which are accelerated in intergalactic or interstellar space through fields of force over enormous distances' §

289. Our kinematic theory in an entirely unforced manner has led to the situation foreseen by Bothe and Kolhorster Without any hypothesis as to gravitation, without the introduction of any arbitrary constants, solely from the hypothesis of the equivalence of the extragalactic nebulae, without having aimed at a theory of the origin of cosmic rays, we have been led to the conclusion that any free particle

† The suggestion that cosmic rays originate from 'super novae' in other galaxies has recently been put forward by Baade and Zwicky

‡ *Zeits fur Phys* , **56**, 777, 1929

§ 'Eine Korpuskularstrahlung konnte ihre Energie in sehr schwachen, dafur aber ungeheuer ausgedehnten Kraftfeldern erlangen, rechnet doch beispielsweise die Entfernung der nichtgalaktischen Nebel nach heutigen Vorstellungen nach Millionen von Lichtjahren '

at large amongst the extra-galactic nebulae attains the velocity of light in a finite time We foresaw the possibility of this conclusion when we considered a single free particle describing its trajectory amongst the nebulae when the latter were idealized to particles The conclusion was then shown to follow rigorously when we contemplated the nebulae in more detail as themselves assemblages of particles in motion

If the primary or pre-primary agent producing the phenomenon of cosmic rays is to be identified with our high-speed particles, we may regard the course of the phenomenon as follows What we mean by a 'particle' we have purposely left vague Our analysis applies strictly to a system of identical particles But clearly in its broad lines it will apply to the motions in a system of unequal particles formed by agglomerative or collision processes out of original, primary, identical particles It would be quite feasible to conduct our analysis for a mixture of two or more sorts of particles, each sort equivalent amongst themselves Thus we may suppose the original system of particles to include electrons, positrons, protons, neutrons, and higher nuclei Some of these interact with one another, producing large aggregates—dust-particles, meteor-particles, cometary particles, stars Once a larger particle has been formed, it will give more 'room' for the other particles left, thus making their further collisions with one another less probable, and at the same time its own increased size will cause a tendency to increase its chance of picking up further particles The net result will be that it will be the smallest of the original, primeval particles which tend to escape collisions After a time we shall have a mixture of smaller particles, charged and uncharged, in free flight, relatively unimpeded in the less congested regions between the galactic nuclei, and all equally subject to the accelerations (which we call gravitational) which we have calculated in the ideal case From time to time the trajectories of such particles will enter congested regions, and certain of them, having hitherto escaped collisions, will enter congested regions at the phase of their trajectories at which they are nearing the speed of light Here they will be likely to undergo collisions—for example with the particles of our own atmosphere The resulting collisions, whether the high-speed particles were charged or uncharged, will be of a drastic character They may well lead directly to atomic disintegrations of the most complicated kind, or they may knock out a positron or other fragment endowed with a

very high energy, which itself might produce disintegration effects Both charged and uncharged particles should be subject equally to gravitational acceleration up to the speed of light, so that we are not bound to identify as the primary or pre-primary agent any particular sort of particle

The observed phenomena the earlier division into primary and secondary radiation, the sudden bursts of atomic nuclei (Hoffmannsche Stosse), the showers of corpuscles (whether these are different or substantially similar phenomena) are all very complicated, and our identification of our high-speed particles with the primary agency is itself sufficiently flexible to afford the possibility of their complications If the observed cosmic-ray phenomenon were too simple a one, our hypothesis would be rendered improbable, for it naturally provides a variety of high-speed missiles each of which should produce its own effects

290. It would be outside the scope of this book to attempt to trace in more detail the course of events following the entry of a high-speed particle into a congested region I content myself with pointing out that the kinematic theory, with its consequence of the existence in every domain of space at every time of some high-speed particles, accounts immediately and simultaneously for the following facts

(1) The corpuscular character of the primary or pre-primary radiation

(2) The high energy of the primary corpuscles An electron possessing a total energy of 10^{11} electron-volts would have a speed of $c(1-1\ 3\times10^{-11})$ cm sec $^{-1}$ (the energy mc^2 of a 'resting electron' is equal to 5×10^5 electron-volts), one possessing an energy of 10^9 electron-volts would have a speed of $c(1-1\ 3\times10^{-7})$ cm sec $^{-1}$ It can be said that the very high energies now found for the primary particles correspond to velocities indistinguishable from that of light

(3) The isotropy of the primary radiation and its independence of sidereal time

291. But our identification abolishes a far more fundamental difficulty which has previously beset attempts at explanation of the cosmic rays This difficulty has been already mentioned above in connexion with the hypotheses of Jeans and Millikan It has been argued that cosmic rays cannot originate in the interior of stars or in the nucleus of an extra-galactic nebula, because they could never get out, even if born there, on the other hand they cannot have been

born in inter-galactic or interstellar space, because the density there is too small to permit the collisions to which the production of the radiations would have to be ascribed. This has seemed an impasse—the 'desperate' position mentioned by Corlin. Our explanation shows that the high-speed corpuscles constituting the primary agent in cosmic ultra-radiation neither come out of the agglomerations which form the nuclei, nor are generated out of rare matter in inter-galactic space. Cosmic rays, or rather the primary agents producing them, are nowhere *born* as such, on our identification. The particles just acquire the high speeds as the inevitable result of continuous acceleration in the 'gravitational' field of the universe. Any particular high-speed particle will possess a value $\mathbf{V}_0$ which identifies it as belonging to some particular sub-system or extra-galactic condensation, but it has not had *to come out of* this particular condensation. It has been on the contrary free ever since the primeval singularity $t = 0$. It may be considered as having equally left any other galactic nucleus, for all are equivalent and were together near $t = 0$. It 'belongs to' a particular sub-system only in a mathematical sense. There is no particular 'here' or 'there' where the particle 'originated'. The singularity $t = 0$ may be supposed to have been equally at any galactic nucleus. Provided the particle in question escapes other collisions, it is freely accelerated from its 'starting' velocity $\mathbf{V}_0$ (possibly at first being decelerated), and this acceleration persists from the dawn of time onwards. The particle only appears as participating in a cosmic-ray phenomenon if it happens to undergo a collision whilst its speed is sufficiently close to that of light. If it does not undergo such a collision, it moves onward through the universe, and we consider its ultimate fate later. It has never had the task of emerging from any particular condensation, other than its passage away from the primeval singularity $t = 0$. Then, and there, it possessed the velocity $\mathbf{V}_0$, as part of that distribution of velocities necessary to conceal the existence, or to display the non-existence, of any preferential velocity-frame. Since $t = 0$ it has been in transit through the universe, attempting centripetally to seek its centre and always failing, until, near the epoch $t = t_l$ in the experience of any chosen observer, it nears the speed of light.

292. We see that the occurrence, in every region of space, at every epoch, of some high-speed particles, is no fortuitous phenomenon,

embroidered on to an otherwise complete world-scheme, but is an essential constituent of the structure of the cosmos, the attaining of the speed of light is an inevitable event in each free particle's history. Cosmic rays are the accompaniment and consequence of this. Cosmic radiation thus appears as an intrinsic phenomenon in the universe, part and parcel of the general scheme of world-structure of which other aspects are the expansion phenomenon, the velocity-distance proportionality, and the subdivision of the material forming the cosmos into nuclear sub-systems.

The existence of the cosmic-ray phenomenon may be looked on as a general confirmation of our scheme. If it had not yet been observed, our kinematic theory would have suggested looking for it. For it is particularly to be noticed that we did not set out to construct an *ad hoc* explanation of cosmic radiation. We have simply made an identification of a predicted phenomenon with an observed one. In this respect our explanation of cosmic radiation, being entirely unforced, is on a different footing from previous attempts at explanation.

293. The general account of cosmic radiation given above includes many observations not directly relevant to our suggestion of the origin of the various phenomena, but I have included it in order that the reader may judge for himself whether there are any observations in direct conflict with the explanation. For my part I cannot see anything inherently improbable in the suggested explanation or any facts which rule it out. On the other hand the suggested explanation covers many of the observed facts, and this affords an *a priori* justification for investigating some of its consequences, which we now proceed to consider.†

Cosmic rays in our own neighbourhood

294. At the epoch t_l of its trajectory (in O's experience) at which a particle attains the velocity c, the direction of its velocity is given by the vector OM_l in the hodograph diagram (Fig. 14), where $|OM_l| = c$. This direction is readily determined from equation (9), Chapter X. If $\mathbf{V}_l$ denotes the velocity-vector here ($|\mathbf{V}_l| = c$), then

$$\mathbf{V}_l = \mathbf{V}_0 - c(1-V_0^2/c^2)^{\frac{1}{2}} \lim_{\xi\to\infty} \frac{\sinh(\eta+\zeta)}{\cosh(\epsilon-\eta-\zeta)}\mathbf{i}$$

$$= \mathbf{V}_0 - c(1-V_0^2/c^2)^{\frac{1}{2}}(\cosh\epsilon+\sinh\epsilon)\mathbf{i},$$

† I am much indebted to Professors R. H. Fowler and P. M. S. Blackett for sympathetic discussion.

or $$\mathbf{V}_l = \mathbf{V}_0-(\mathbf{V}_0\cdot\mathbf{i})\mathbf{i}-c[1-(\mathbf{V}_0^2/c^2)+(\mathbf{V}_0\cdot\mathbf{i})^2/c^2]^{\frac{1}{2}}\mathbf{i} \tag{6}$$
This satisfies identically $|\mathbf{V}_l| = c$, as is readily verified.

Now let us examine those particles which arrive in our own neighbourhood (neighbourhood of O) with speeds approaching c. For simplicity we shall consider our own neighbourhood to be the centre of our own galaxy.† The position of such particles is therefore given by $\mathbf{P} = \mathbf{O} = 0$. Putting $\mathbf{P} = 0$ in (10) and (11), Chapter X, and putting $t_l = t$, our present epoch, we have

$$0 = \mathbf{V}_0 t-cX^{\frac{1}{2}}\frac{\sinh\eta_l}{\cosh\epsilon}\mathbf{i}, \tag{7}$$

where $$t = \frac{X^{\frac{1}{2}}}{(1-V_0^2/c^2)^{\frac{1}{2}}}\frac{\cosh(\epsilon-\eta_l)}{\cosh\epsilon}. \tag{8}$$

Eliminating t, $$\frac{\mathbf{V}_0/c}{(1-V_0^2/c^2)^{\frac{1}{2}}} = \frac{\sinh\eta_l}{\cosh(\epsilon-\eta_l)}\mathbf{i}. \tag{9}$$

Hence $\mathbf{V}_0$ is parallel to $\mathbf{i}$. Hence by (6), $\mathbf{V}_l$ is also parallel to $\mathbf{i}$. Multiplying (9) scalarly by $\mathbf{i}$, we have

$$\sinh\epsilon = (\sinh\eta_l)/\cosh(\epsilon-\eta_l),$$

which gives at once $$\epsilon = \eta_l. \tag{10}$$

Hence, by (9), $$\mathbf{V}_0 = c\tanh\eta_l\,\mathbf{i}, \tag{11}$$

and then, from (6), $$\mathbf{V}_l = -c\mathbf{i}. \tag{12}$$

295. It follows that particles now arriving in our vicinity with the velocity of light, in the direction $-\mathbf{i}$, had at $t = 0$ the velocity $\mathbf{V}_0 = c\tanh\eta_l\,\mathbf{i}$. Hence the present position (at time t) of the nucleus of the sub-system to which these particles 'belong', is

$$\mathbf{P}_0 = ct\tanh\eta_l\,\mathbf{i}. \tag{13}$$

Since $\tanh\eta_l < 1$, P_0 is inside the expanding light-sphere $|\mathbf{P}| = ct$, and the distance of the *apparent* source of the swift particles is just $ct\tanh\eta_l$. We recall that by (4) the intensity of radiation of such particles *of all velocities* is

$$\frac{\pi C\eta_l}{c^2t^3} \tag{14}$$

per cm.² per second. This connects the intensity with the distance of the apparent source (though it may equally be supposed that the particles 'originated' at ourselves or at any other nucleus). Since t is

† The immediately following analysis can be readily adapted to discuss particles arriving *near* the centre of a galaxy, with speeds *near* c.

about 0 6 $\times 10^{17}$ seconds, a rate of arrival of one particle per cm ^{2} per sec would give $C\eta_l \sim 6 \times 10^{70}$ The actual value of $C\eta_l$ or

$$\frac{1}{2}\int\limits_1^\infty \frac{\xi-1}{\xi^{\frac{1}{2}}}\psi(\xi)\,d\xi$$

must be much larger than this, as only a fraction of the number $\pi C\eta_l/c^2t^3$ will represent high-speed particles This is sufficient to show that the product of the two pure numbers C and η_l (both world-constants) is a very large number A further check is provided by formula (5), Chapter X It is clear from this formula that away from the singularities $\xi = 1$, i e away from the nebular nuclei, the total number of particles per cm ^{3}, of all velocities, is comparable with $2 \times 2\pi C\eta_l/c^3t^3$, though it is not exactly equal to this This must be less than the mean density of matter in space obtained by smoothing out all the galaxies, say less than 10^{-30} gramme cm ^{3} Identifying the primary particles, for the sake of argument, with protons, we have

$$\frac{4\pi C\eta_l}{c^3t^3} \lesssim \frac{10^{-30}}{1{\cdot}65\times 10^{-24}}$$

or

$$C\eta_l \lesssim 3\times 10^{74},$$

where the sign $\lesssim$ means that $C\eta_l$ will not exceed but may be of the order of 3×10^{74}

More delicate analysis, of accelerations, etc , which I do not here attempt, would be necessary to separate the product $C\eta_l$ into its factors, but the evidence suggests that η_l is probably a large number In that case $\tanh\eta_l$ will approach unity in magnitude, and cosmic-ray particles therefore 'belong' to nuclear sub-systems near though inside the expanding frontier of the universe This conclusion, though probable, is only conjectural as far as we have gone Careful discussion of accelerations near and outside a galactic system should serve to determine the ratio $C/(\xi-1)^{\frac{3}{2}}\psi(\xi)$ for $\xi \sim 1$, which is probably of the order of unity, and since $\psi(\xi)$ affords a measure of the density, C is in principle determinable

296. The history of every particle ultimately taking part in a cosmic-ray phenomenon near ourselves is as follows The particle may be considered as having left our own neighbourhood at $t = 0$ with velocity $c\tanh\eta_l\,\mathbf{i}$ (numerically probably not much less than c) in an outward direction, it is retarded in the direction of $\mathbf{i}$ until it

comes to rest relative to our own system, then falls back and attains the speed c in the direction $-\mathbf{1}$ as it returns to our neighbourhood If the reader takes the trouble to work out the hodograph relations for such a particle, and the behaviour of the apparent centre of the universe to the particle-observer moving with it, he will find that the particle is always ahead of the apparent centre in the direction $\mathbf{1}$, which accounts for the apparent attraction towards O At time t itself, in the experience of O, the particle-observer moving with the swift particle assigns to the apparent centre (to him) of the universe, reckoned from O, the position $-ct\,\mathbf{1}$

297. It should be noticed that the swift corpuscles incident on the earth will not consist solely of 'primary' particles which have arrived in our neighbourhood with speeds approaching that of light, they will include other high-speed particles which have originated from collisions undergone by 'primary' high-speed particles in other regions of the universe Such other particles, liberated in this way, will themselves become free particles acted on by the 'gravitational field' of the universe, and will undergo accelerations and pursue trajectories of the type already discussed They may therefore in turn reach our neighbourhood with speeds approaching that of light This of itself suggests that the observed cosmic-ray phenomenon should be extremely complicated The calculations of the present section relate simply to the arrival of the primary particles in our own neighbourhood Secondary particles should equally arrive from all directions in space

Evolutionary consequences

298 Several other deductions of cosmic interest may be made from (14) As t increases, the intensity of primary cosmic radiation should diminish It should be diminishing at the present epoch at the rate of 3 parts in $0\,6 \times 10^{17}$ per second, or 1 part in 7×10^{8} per year Conversely, in past ages it must have been more intense, theoretically of very great intensity near 'creation', $t \sim 0$ This opens up enormous vistas of speculation Conditions on the earth may have been totally different, due to the enhanced intensity of cosmic radiation, 10^{9} years ago, and the influences affecting living organisms in early geological time must have included the effects of strong cosmic radiation. Even the problem of the origin of life may take a radically different complexion if it could be assured that cosmic radiation was then potent

Knowing as we do that living tissues are markedly affected by the passage through them of rays from radioactive substances—that prolixity of cell-growth on the one hand or retardation of cell-growth on the other hand may result—we must recognize the possibility that metabolism and cell-division may have proceeded at totally different rates in earlier epochs. The whole course of evolution—the non-repeating development of particular floras like the carboniferous, the extinction of species, the emergence of higher forms, and incidence of cancer mortality—may have been conditioned by the steady diminution during evolutionary time of the intensity of ultra-radiation. The mere energy, apart from nuclear and atomic disintegration effects, of a radiation which even now is comparable with star-light (Regener) may have had enormous effects on terrestrial conditions. We may conjecture that the intensity of cosmic ultra-radiation may have always been comparable with the total radiation of energy in the forms of heat and light from the stars and nebulae, for both must have then been greater in former ages, when the nebulae were smaller and closer together.

Still further inferences may be made from (14). This formula holds good for any region of the universe if t is the local time. Hence regions at great distances, approaching ct from us in our time, must be exposed at our present epoch of observation to much greater intensity of primary cosmic radiation than we are exposed to. They, the nebulae at great distances, in our 'now' and under our observation, are experiencing the conditions of the *early* history of our particular portion of the cosmos. Given a large enough telescope, the apparent congestion at great distances which our analysis predicts would be accompanied in our vision with strong cosmic radiation, so that even now we can in principle observe portions of the universe where cosmic ultra-radiation is all-powerful.

XIII

COSMIC CLOUDS

Motion following $|\mathbf{V}| = c$

299 WE now investigate the trajectory of a particle after the epoch at which it attains the velocity of light There are *a priori* two possibilities Either (*a*) it continues to move with the speed of light, or (*b*) its speed diminishes The reader may object that there is a third possibility, namely that the speed increases above that of light The universe would then include, in every neighbourhood at every epoch, particles moving in all directions with speeds exceeding that of light, and such particles, when directly receding from the observer, would be unobservable This would be equivalent to the conclusion that the totality of things would contain unobservable particles, which is an unverifiable proposition Further, the passage through and beyond the speed of light for such particles would be equivalent to annihilation of matter and energy (the reverse of the creation of matter), which is not observed All our kinematic formulae in their present forms become meaningless for $|\mathbf{V}| > c$, and the attaining of speeds exceeding c is incompatible with the whole of existing dynamics and electromagnetism I hesitate to believe that a rational scheme of things could be arranged containing a description of such particles, and I leave it to others to contemplate this possibility and work out its formulation if it be possible It seems more reasonable to confine attention to the two possibilities enumerated above

300 I have elsewhere† investigated the behaviour of a system of particles satisfying the cosmological principle and possessing the speed of light The analysis given in this book has to be reconsidered, in large part, for particles such that $|\mathbf{V}| = c$, but it is possible to find the limit of the acceleration and the nature of the distribution In the memoir cited, I showed that such a system is a possible system, but that the 'particles', to be described fully, must each be characterized by a new parameter ν, possessing all the properties of a frequency I showed that such particles would be propagated in plane waves and I obtained a formula giving their distribution Such a system of particles moving with the speed c then appeared to be indistinguishable from a system of photons, arranged so as to con-

† *Zeits fur Astrophys*, **6**, 83, Part III, 1933

stitute plane waves of light moving in all directions There thus appeared a complete disjunction between c-moving particles and particles with other speeds, corresponding to the observed disjunction between matter and light The photons were essentially a different kind of entity from the abstract particles of our present general theory

301. The particles discussed in the foregoing chapters have been *material* particles This identification is certainly necessary if we are to attribute to such particles, when moving with nearly the speed of light, the effects associated with cosmic radiation Further, our identification of the regions of congestion of the sub-systems of such particles with the nuclei of extra-galactic nebulae demands also that the particles we are discussing be material particles

302. Now if such a material particle, after attaining the speed of light, continued to move with the same speed, it would constitute a photon of infinite energy, or infinite frequency (to employ the language of dynamics), and a catastrophe would result To avoid this, we are driven to suppose that the velocity only momentarily reaches c, and that it then decreases again, this being alternative (b) Alternative (a) would mean that in effect the material particle turned suddenly into light for no particular reason—a hypothesis too wild to entertain Alternative (b) preserves the disjunction between matter and light, a material particle remains a material particle For whilst a material particle *might* happen to undergo a collision at the exact moment at which its velocity was c, and so give rise to a cosmic ray which would disrupt the universe, the probability of such an occurrence has been seen to be zero, for, as we saw, the number of particles arriving in a given domain, or impingeing on a given area, with the precise velocity c, is vanishingly small, by formulae (5) and (6), Chapter X

Fortunately no specific assumption on our part is required Our mathematics provides of its own accord for the decrease of velocity after $|\mathbf{V}| = c$ Any other behaviour would indeed be incompatible with our analysis

303 We know (equation (7′), Chapter VIII, on using (21), Chapter IX) that

$$\frac{d\xi}{dt} = +2\frac{Z}{X}\frac{C}{(\xi-1)^{\frac{1}{2}}\psi(\xi)}, \qquad (C > 0) \tag{1}$$

Since $\int^{\xi} (\xi-1)^{\frac{1}{2}}\psi(\xi)\, d\xi$ has been shown to converge as $\xi \to \infty$, it follows that $(\xi-1)^{\frac{1}{2}}\psi(\xi) \to 0$ as $\xi \to \infty$, and tends to zero in fact faster than

N/ξ, where N is any constant Hence by (1), since Z and X correspond to a finite position of P, when $|\mathbf{V}| = c$, we see that $d\xi/dt \to \infty$ as $\xi \to \infty$ This is otherwise obvious, since $\xi \to \infty$ as $t \to t_l$ Since $(\xi-1)^{\frac{1}{2}}\psi(\xi)$ passes through a zero as $\xi \to \infty$, and since $\psi(\xi)$ is always positive (it defines a number of particles), $(\xi-1)^{\frac{1}{2}}\psi(\xi)$ may be taken to change sign as ξ passes through ∞, the change of sign being automatically provided by the square root In that case, by (1), ξ begins to decrease again for $t > t_l$

304. We are here deriving a physical consequence from a mathematical situation The mathematical situation is that a square root is two-valued, and the mathematics inevitably invites consideration of the second, negative value, even though we originally selected the positive value Such making of a physical inference on grounds of mathematical form is very common, especially in modern mathematical physics For instance Dirac inferred the existence of particles in states of negative energy simply from the quadratic form of the relativistic Hamiltonian in the theory of the motion of a particle, and the existence of polarization-charge in a vacuum has also been inferred similarly In such cases we are trusting the mathematics to look after the physical situation If my procedure here is criticized, I tell my critics to go and criticize the similar procedure of Dirac and others and to examine the fruitfulness of their procedures A critic must be consistent, and I see no reason why I should defend a procedure analogous to similar procedures used in the quantum theory until the latter procedures are successfully criticized The justification of the procedure ultimately rests on its success, and I propose to follow out its consequences We shall find that the resulting analysis goes far towards an understanding of the history of the universe, and explains many of the observed features of the universe

305. Since ξ is Z^2/XY and Z and X are finite and non-zero at $t = t_l$, Y must increase from zero as ξ decreases from ∞ Hence $|\mathbf{V}|$ decreases

The square root $(\xi-1)^{\frac{1}{2}}$ in (1) arises from the term

$$\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}$$

in the acceleration formula If $(\xi-1)^{\frac{1}{2}}\psi(\xi)$ changes sign as it passes through its zero, so must $(\xi-1)^{\frac{3}{2}}\psi(\xi)$ change sign Hence the second term in the acceleration formula changes sign as ξ increases to ∞ and decreases again

Now as we have seen, the acceleration tends to zero as $\xi \to \infty$ We now notice that the second term in the acceleration formula is dominant to the first as ξ approaches ∞ For the formula may be written

$$-(\mathbf{P}-\mathbf{V}t)\frac{Z^2}{X^2}\left[\frac{1}{\xi}+\frac{C}{\xi(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right] \tag{2}$$

Now since $\int^{\xi} (\xi-1)^{\frac{1}{2}}\psi(\xi)\, d\xi$ converges as $\xi \to \infty$, $(\xi-1)^{\frac{1}{2}}\psi(\xi)$ is less than δ/ξ, where δ is any constant, and so $\xi^{-1}(\xi-1)^{-\frac{3}{2}}[\psi(\xi)]^{-1}$ is greater than $1/\delta\xi$ for ξ large Hence the coefficient of $(\mathbf{P}-\mathbf{V}t)$ passes through a zero and changes sign as ξ reaches ∞ and decreases again

306. To say that in the fundamental acceleration formula

$$\frac{d\mathbf{V}}{dt} = -(\mathbf{P}-\mathbf{V}t)\frac{Y}{X}\left[1+\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right] \tag{2'}$$

$C/(\xi-1)^{\frac{3}{2}}\psi(\xi)$ changes sign as ξ passes through ∞ is equivalent to asserting that a new domain of motions is prescribed for $\infty > \xi > 1$ We have carried out integrations for $C > 0, 1 \leqslant \xi < \infty, (\xi-1)^{\frac{1}{2}}\psi(\xi) > 0$, we now desire to carry out integrations for $C > 0$, $\infty > \xi \geqslant 1$,† $(\xi-1)^{\frac{1}{2}}\psi(\xi) < 0$ We originally took the constant of integration C positive to ensure $X \to 0$ as $\xi \to 1$ $(t \to 0)$, we could equally have carried out our integrations for $C < 0$, which would be fully equivalent to reversing the sign of the square root in the X-integral We should then have had $X \to \infty$ as $\xi \to 1$ (ξ decreasing), and as we shall see in a moment this would correspond to $t \to \infty$ We should simply have been beginning our integration at the wrong end! We now see that the form of the analysis compels us in effect to consider the consequences of taking C negative

The term $C/(\xi-1)^{\frac{3}{2}}\psi(\xi)$ in the acceleration formula arose from the differential equation (4) of Chapter IX, a consequence of Boltzmann's equation We were compelled to select a sign for $(\xi-1)^{\frac{3}{2}}$ in order to be definite—it did not matter which Choosing the positive sign, the constant C, for $\xi \sim 1, t \sim 0$, had to be positive (This gave us the sign of the acceleration for resting particles near the observer) We now see that the mathematics insists on our considering the circumstances in which the square root has the opposite sign This we should have ultimately had to do in any case, for mathematical completeness The square root $(\xi-1)^{\frac{1}{2}}$ turned up near the outset of our investigations, as soon as we imposed the condition of conservation of particle-

† For the meaning of this notation, see footnote on p 150

number This square root was no accidental concomitant of the analysis The analysis, as it were, foresaw from the start the very eventuality we are now discussing and provided for the eventuality beforehand The eventuality arose during physical considerations, the speed of the particle rose to that of light The sequel to the eventuality is taken care of by the mathematics, which has throughout introduced in advance square roots whose signs could change

I know of no counterpart to this situation in *current* relativistic cosmology The latter has not pushed so far as we have done the discussion of the trajectory of a free particle But possibly in relativistic cosmology a geodesic becomes tangent somewhere to a nul-geodesic Similar consequences would then follow

307. Hitherto we have been able to give immediate physical interpretations of all our formulae We saw, for example, that the acceleration formula could be interpreted as giving a pull on the free particle directed towards the apparent centre of the universe in the frame in which the free particle was momentarily at rest If we seek a physical interpretation of the 'change of sign' phenomenon, we must now say that the free particle is repelled from the apparent centre of the universe Gravitation, *as it were*, changes sign The result is the converse of the situation previously described In spite of being repelled from the apparent centre, the free particle now gets nearer and nearer to it, and the particle may be expected ultimately to fall into the 'centre' of the universe We shall investigate in due course whether this happens. We could put it better by saying that the centre of the universe overtakes the particle In the first phase of the motion ($\xi < \infty, t < t_l$) the free particle strives to fall into the apparent centre, but gets farther and farther away from it, in the second phase of the motion ($\infty > \xi, t > t_l$) the free particle strives to get away from the apparent centre, but actually approaches it The apparent paradoxes are resolved when we remember that the 'apparent centre' is not some particular particle, but a kinematically defined centre, analogous to the instantaneous centre of rotation of a rigid lamina, which has its own locus in the rigid lamina In our case the instantaneous apparent centre has its own locus in the universe and the net result is that the particle, which coincides with the apparent centre at $t = 0$ but is projected with $V = V_0$ from it, ultimately arrives at (or, as we shall see, near) the apparent centre (to itself) as $t \to \infty$

There is no actual change of sign of gravitation, only a change of sign in the particular formulae we have used to describe the motion, formulae which give the assignment of epochs, distances, and velocities in the experience of any fundamental observer O To a particle-observer moving with the free particle the particle appears first to recede from the apparent centre, then to catch it up, in other words the apparent centre recedes until it reaches the confines of the expanding system, then approaches Nothing *happens* to the particle as it passes through $|V| = c$, that event, in O's experience, simply marks, in the free particle's experience, the epoch at which the apparent centre reaches the confines of the observable system, than which it can go no farther

We see again how contradictory it would be to suppose that the free particle passed to velocities exceeding c, for then the apparent centre would pass outside what to O is the boundary of the system, and a particle inside the region of observable events would be pointing to a hypothetical unobservable point as the apparent centre Any attempt, in fact, to describe the appearance of the system to a particle-observer whose velocity passed to above c would break down in a maze of contradictions with experience—he would see material being annihilated in all directions

The analytical behaviour, a change of sign of a square root, is thus merely an accidental feature of the method of description used by a fundamental observer We have chosen his as the eyes through which we view the system Some eye there must be, for the system to be described at all When the eye is that of the observer on the free particle, he simply sees himself pursuing the apparent centre (though initially projected away from it) and ultimately coming into its neighbourhood

I admit that all this is surprising It is surprising because our usual modes of visualizing physical processes are neither powerful nor reliable The mathematics here serves us better But guided by it, we are led to a rational, if unexpected, description of the course of events

Integration of trajectories for $t > t_l$, $\infty > \xi$

308 We now integrate the equations of motion

$$\frac{d\mathbf{P}}{dt} = \mathbf{V}, \qquad \frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}\left[-1+\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right], \tag{3}$$

which give the continuations of the trajectories after the particle

concerned has attained the speed of light The integrations can be carried out as before All we have to do is to write down our previous integrals with the minus sign placed before the symbol C Six new constants of integration appear For any given particle of a trajectory for $\infty > \xi$, these can be chosen so as to fit the arc continuously on to the preceding arc of the same trajectory, along which $\xi < \infty$, but it is convenient to consider first the integrals pure and simple, and carry out the fitting of the later arcs ($\infty > \xi$) on to the earlier arcs ($\xi < \infty$) afterwards The 6 new constants of integration can be interpreted physically later, for the moment we regard them merely as constants thrown up by the course of the integration

For convenience I shall employ the symbol η to denote the same function of the variable ξ as before, but a prime ($'$) will be used to denote other *corresponding* functions of ξ on the arc $\infty > \xi$ (ξ decreasing) The notation will be clear as it develops The symbol C will continue to denote the same positive number as before

309. The X-integral now takes the form

$$X' = X_2 \exp\left[-\frac{1}{C}\int\limits_{\xi_2}^{\xi} (\xi-1)^{\frac{1}{2}}\psi(\xi)\, d\xi\right], \tag{4}$$

where X_2 is the value taken by X (now called X', where

$$X' = t'^2 - \mathbf{P}'^2/c^2,$$

t', $\mathbf{P}'$ being the epoch and position along the arc $\infty > \xi$) at the arbitrary value $\xi = \xi_2$ As ξ steadily decreases from $\xi = \infty$, X' steadily increases, and since the integral in (4) diverges as $\xi \to 1$, $X' \to \infty$ as $\xi \to 1$ (ξ decreasing)

Since at $\xi = \infty$, $X = X_l$, $X' = X_l'$, for one and the same trajectory, we have by continuity

$$X_l = X_l', \tag{5}$$

where

$$X_l = X_1 \exp\left[\frac{1}{C}\int\limits_{\xi_1}^{\infty} (\xi-1)^{\frac{1}{2}}\psi(\xi)\, d\xi\right],$$

$$X_l' = X_2 \exp\left[-\frac{1}{C}\int\limits_{\xi_2}^{\infty} (\xi-1)^{\frac{1}{2}}\psi(\xi)\, d\xi\right]$$

Since, from previous work,

$$X = X_1 \exp\left[\frac{1}{C}\int\limits_{\xi_1}^{\xi} (\xi-1)^{\frac{1}{2}}\psi(\xi)\, d\xi)\right],$$

we have

$$X = X_l \exp\left[-\frac{1}{C}\int_{\xi}^{\infty} (\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right] \qquad (\xi \text{ increasing}), \qquad (6)$$

$$X' = X_l' \exp\left[\frac{1}{C}\int_{\xi}^{\infty} (\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right] \qquad (\xi \text{ decreasing}) \qquad (6')$$

Hence for the same value of ξ on the two arcs (ξ increasing and ξ decreasing) we have

$$X(\xi)X'(\xi) = X_l X_l' = X_l^2 \qquad (7)$$

310 As we are denoting by the same symbol η the function

$$\eta(\xi) = \frac{1}{2C}\int_{1}^{\xi} \frac{\xi-1}{\xi^{\frac{1}{2}}}\psi(\xi)\,d\xi, \qquad (8)$$

all we have to do to obtain our new integrals is to reverse the sign of η in the previous integrals and prime all other symbols except ζ, which continues to be defined by

$$\cosh\zeta = +\xi^{\frac{1}{2}}, \qquad \sinh\zeta = +(\xi-1)^{\frac{1}{2}}$$

I must leave to the reader the verification that the plus sign has to be retained in the merely conventional definition of the symbol ζ

311. We now tabulate the corresponding integrals for ξ increasing and ξ decreasing in parallel columns

ξ increasing $(1 < \xi < \infty)$	*ξ decreasing* $(\infty > \xi \geqslant 1)$	
$\sinh\epsilon = \dfrac{(\mathbf{V}_0 \cdot \mathbf{i})/c}{(1-V_0^2/c^2)^{\frac{1}{2}}}$	$\sinh\epsilon' = \dfrac{(\mathbf{V}_0' \cdot \mathbf{i}')/c}{(1-V_0'^2/c^2)^{\frac{1}{2}}}$	(9)
$\mathbf{V} = \mathbf{V}_0 - c(1-V_0^2/c^2)^{\frac{1}{2}}\dfrac{\sinh(\eta+\zeta)}{\cosh(\epsilon-\eta-\zeta)}\mathbf{i}$	$\mathbf{V}' = \mathbf{V}_0' - c(1-V_0'^2/c^2)\dfrac{\sinh(-\eta+\zeta)}{\cosh(\epsilon'+\eta-\zeta)}\mathbf{i}'$	(10)
$\mathbf{P} = \mathbf{V}_0 t - cX^{\frac{1}{2}}\dfrac{\sinh\eta}{\cosh\epsilon}\mathbf{i}$	$\mathbf{P}' = \mathbf{V}_0' t' + cX'^{\frac{1}{2}}\dfrac{\sinh\eta}{\cosh\epsilon'}\mathbf{i}'$	(11)
$t = \dfrac{X^{\frac{1}{2}}}{(1-V_0^2/c^2)^{\frac{1}{2}}}\dfrac{\cosh(\epsilon-\eta)}{\cosh\epsilon}$	$t' = \dfrac{X'^{\frac{1}{2}}}{(1-V_0'^2/c^2)^{\frac{1}{2}}}\dfrac{\cosh(\epsilon'+\eta)}{\cosh\epsilon'}$	(12)
$\mathbf{P}-\mathbf{V}t = cX^{\frac{1}{2}}\dfrac{\sinh\zeta}{\cosh(\epsilon-\eta-\zeta)}\mathbf{i}$	$\mathbf{P}'-\mathbf{V}'t' = cX'^{\frac{1}{2}}\dfrac{\sinh\zeta}{\cosh(\epsilon'+\eta-\zeta)}\mathbf{i}'$	(13)
$Y^{\frac{1}{2}} = (1-V_0^2/c^2)^{\frac{1}{2}}\dfrac{\cosh\epsilon}{\cosh(\epsilon-\eta-\zeta)}$	$Y'^{\frac{1}{2}} = (1-V_0'^2/c^2)^{\frac{1}{2}}\dfrac{\cosh\epsilon'}{\cosh(\epsilon'+\eta-\zeta)}$	(14)
$X = X_l \exp\left[-\dfrac{1}{C}\displaystyle\int_{\xi}^{\infty}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right] \equiv t^2-\mathbf{P}^2/c^2$	$X' = X_l' \exp\left[+\dfrac{1}{C}\displaystyle\int_{\xi}^{\infty}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right] \equiv t'^2-\mathbf{P}'^2/c^2$	(15)
$\xi = \dfrac{(t-\mathbf{P}\cdot\mathbf{V}/c^2)^2}{\{t^2-(\mathbf{P}^2/c^2)\}\{1-(\mathbf{V}^2/c^2)\}}$	$\xi = \dfrac{(t'-\mathbf{P}'\cdot\mathbf{V}'/c^2)^2}{\{t'^2-(\mathbf{P}'^2/c^2)\}\{1-(\mathbf{V}'^2/c^2)\}}$	(16)

In the above, $\mathbf{V}_0'$, $\mathbf{i}'$, and X_l' provide 6 new arbitrary constants of integration On any given complete trajectory, the values of $\mathbf{V}_0', \mathbf{i}', X_l'$ are determinate in terms of those of $\mathbf{V}_0, \mathbf{i}, X_l$ from the conditions of continuity at the place where $\xi = \infty$, $|\mathbf{V}| = c$ These relations we proceed to find

Relations between constants

312 At the point of a trajectory where $\xi = \infty$, $|\mathbf{V}| = c$, $\eta = \eta_l$, the two vectors $\mathbf{V}$, $\mathbf{V}'$ are equal, the two vectors $\mathbf{P}, \mathbf{P}'$ are equal, and $t = t'$. This appears to give 7 conditions but they reduce to 6, since the relation $|\mathbf{V}| = |\mathbf{V}'| = c$ is identically satisfied at $\xi = \infty$

The first condition of fit, namely $X_l = X_l'$, we have already obtained It follows also from (15) since X and X' are continuous at $\xi = \infty$

Since the values of $\mathbf{P}-\mathbf{V}t$, $\mathbf{P}'-\mathbf{V}'t'$ are continuous at $\xi = \infty$, or $\zeta = \infty$, we have at once from (13)

$$\mathbf{i} = \mathbf{i}', \tag{17}$$

and
$$\epsilon-\eta_l = \epsilon'+\eta_l \tag{18}$$

The latter gives
$$\epsilon-\epsilon' = 2\eta_l \tag{18'}$$

313 Equating the values $\mathbf{V}_l, \mathbf{V}_l'$ of $\mathbf{V}$ and $\mathbf{V}'$ at $\xi = \infty$, and multiplying vectorially by $\mathbf{i}$ (or $\mathbf{i}'$) we have

$$\mathbf{V}_0 \wedge \mathbf{i} = \mathbf{V}_0' \wedge \mathbf{i}' \tag{19}$$

Hence the plane of $\mathbf{V}_0$ and $\mathbf{i}$ coincides with the plane of $\mathbf{V}_0'$ and $\mathbf{i}'$ Since the first branch of the trajectory lies entirely in the plane of $\mathbf{V}_0$ and $\mathbf{i}$, and the second branch in that of $\mathbf{V}_0'$ and $\mathbf{i}'$, it follows that the two branches of the trajectory are coplanar.

314. We now write down the equality $\mathbf{P}/t = \mathbf{P}'/t$ at $\xi = \infty$, by equating the quotients of expressions (11), (12) with $\eta = \eta_l$, $X = X' = X_l$ We then carry out the following three different transformations of the resulting equality

(i) insert for t_l and t_l' their expressions as given by (12) respectively,

(ii) insert for both t_l and t_l' the *first* of expressions (12),

(iii) insert for both t_l and t_l' the *second* of expressions (12)

We obtain the following three relations, after some use of (18),

$$\mathbf{V}_0 - c\mathbf{i}(1-V_0^2/c^2)^{\frac{1}{2}}\frac{\sinh\eta_l}{\cosh(\epsilon-\eta_l)} = \mathbf{V}_0' + c\mathbf{i}(1-V_0'^2/c^2)^{\frac{1}{2}}\frac{\sinh\eta_l}{\cosh(\epsilon'+\eta_l)}, \tag{20}$$

$$\mathbf{V}_0' = \mathbf{V}_0 - c\mathbf{i}(1-V_0^2/c^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)}, \tag{21}$$

$$\mathbf{V}_0 = \mathbf{V}_0' + c\mathbf{i}(1-V_0'^2/c^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(\epsilon'+2\eta_l)} \tag{22}$$

These three relations between the vectors $\mathbf{V}_0$ and $\mathbf{V}_0'$ are compatible, taking into account (9), and each implies the other two. We notice that these imply

$$(1 - V_0^2/c^2)^{\frac{1}{2}}\cosh\epsilon = (1 - V_0'^2/c^2)^{\frac{1}{2}}\cosh\epsilon', \tag{23}$$

a useful relation which is, however, already implied by (19).

315. We have now connected the 6 new constants with the 6 old ones. X_l', $\mathbf{i}'$ are equal respectively to X_l, $\mathbf{i}$. Relation (18) determines ϵ' in terms of ϵ, given the world constant η_l, and (21) determines $\mathbf{V}_0'$ in terms of $\mathbf{V}_0$ and η_l. Relation (18) is extremely beautiful, and its simplicity shows that we made a wise choice of notation in introducing the combination of $\mathbf{V}_0$ and $\mathbf{i}$ which we call ϵ. The reciprocal character of relations (20), (21), and (22), which we shall discuss shortly, should also be noticed.

316. Since any one of the three relations (20), (21), (22) is equivalent to 3 scalar relations, and since $\mathbf{i}$ is equivalent to only 2 scalar numbers, there must be one scalar relation between $\mathbf{V}_0$ and $\mathbf{V}_0'$ independent of $\mathbf{i}$. This is found to be†

$$\tanh^2 2\eta_l = \frac{c^2(\mathbf{V}_0' - \mathbf{V}_0)^2 - (\mathbf{V}_0' \times \mathbf{V}_0)^2}{(c^2 - \mathbf{V}_0\cdot\mathbf{V}_0')^2}. \tag{24}$$

317. With the aid of the above relations, we can express the formulae in the second column of integrals (9)–(14) in terms of the original constants of integration $\mathbf{V}_0$, $\mathbf{i}$, X_l. Some calculations whose details we omit give

$$\mathbf{V}' = \mathbf{V}_0 - c(1 - V_0^2/c^2)^{\frac{1}{2}}\frac{\sinh(2\eta_l - \eta + \zeta)}{\cosh(\epsilon - 2\eta_l + \eta - \zeta)}\mathbf{i}, \tag{10'}$$

$$\mathbf{P}' = \mathbf{V}_0 t' - cX_l X^{-\frac{1}{2}}\frac{\sinh(2\eta_l - \eta)}{\cosh\epsilon}\mathbf{i}, \tag{11'}$$

$$t' = \frac{X_l X^{-\frac{1}{2}}}{(1 - V_0^2/c^2)^{\frac{1}{2}}}\frac{\cosh(\epsilon - 2\eta_l + \eta)}{\cosh\epsilon}, \tag{12'}$$

$$\mathbf{P}' - \mathbf{V}'t' = cX_l X^{-\frac{1}{2}}\frac{\sinh\zeta}{\cosh(\epsilon - 2\eta_l + \eta - \zeta)}\mathbf{i}, \tag{13'}$$

$$\Gamma'^{\frac{1}{2}} = (1 - V_0^2/c^2)^{\frac{1}{2}}\frac{\cosh\epsilon}{\cosh(\epsilon - 2\eta_l + \eta - \zeta)}. \tag{14'}$$

In these formulae, X is the function $X(\xi)$ of ξ defined by (6). These formulae exhibit explicitly the continuity, at $\epsilon = \epsilon'$, $\eta = \eta_l$, $X = X_l$,

† See Note 6.

of the two branches of the trajectory For they have been obtained by transformation of the second column of integrals (9)–(14), and they reduce to the first column on putting $\eta = \eta_l$, $X = X_l$, and letting $\zeta \to \infty$

318 We now comment on a remarkable point We have considered the complete family of integrals of the two separate sets of differential equations,

$$\frac{d\mathbf{P}}{dt} = \mathbf{V}, \qquad \frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}\left(-1-\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right), \tag{25}$$

$$\frac{d\mathbf{P}'}{dt'} = \mathbf{V}', \qquad \frac{d\mathbf{V}'}{dt'} = (\mathbf{P}'-\mathbf{V}'t')\frac{Y'}{X'}\left[-1+\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right] \tag{25'}$$

It follows that if we are given only the circumstances of projection $\mathbf{P}, \mathbf{V}, t$ the acceleration appears to be two-valued But actually the complete family of trajectories is determinate This is because we are confined to real trajectories, i e trajectories starting at $t = 0$, $\mathbf{P} = 0$ Along the branch distinguished by the primed symbols, $d\xi/dt$ is negative, along the other branch, $d\xi/dt$ is positive If then decreasing values of t are considered, ξ is increasing along the primed branch, and then along this branch, by the second of (12), t' does not tend to zero but to a constant value t_l as $\xi \to \infty$ and the zero of time is not accessible along this branch Changing over to the unprimed branch, we find that t decreases to zero as required, and we follow backwards a trajectory already enumerated The complete set of constants $\mathbf{V}_0, \mathbf{i}, X_l$ or $\mathbf{V}_0', \mathbf{i}', X_l'$ enumerates completely all trajectories that are observable from $t = 0$ onwards

We can put the matter another way Suppose that we fix the circumstances of projection, $\mathbf{P}, \mathbf{V}, t$, and then attempt to solve the equations for the corresponding constants of integration, namely $\mathbf{V}_0, \mathbf{i}, X_l$, from the integrals of (25) and $\mathbf{V}_0', \mathbf{i}', X_l'$ from the integrals of (25′), treating these as independent constants of integration not necessarily connected by the continuity relations (5), (17), and (20). Then we find incompatibility unless $\xi = \infty$, $|\mathbf{V}| = c$ We can only solve either for the one set or the other An unprimed branch only coincides in epoch, position, and velocity with a primed branch at the point $\xi = \infty$, $|\mathbf{V}| = c$, where they are branches of *one and the same* trajectory For the complete family of trajectories issuing from $t = 0$, $\mathbf{P} = 0$, with the totality of corresponding accessible sets of values of $t, \mathbf{P}, \mathbf{V}$, the acceleration is always completely determinate

We can put the matter still another way The equations of continuity $\mathbf{P} = \mathbf{P}'$, $\mathbf{V} = \mathbf{V}'$, $t = t'$ provide 7 conditions to be satisfied, whilst we have only 6 constants to determine They are only compatible when $\xi = \infty$ The fact should be remembered that ξ is not a mere parameter, but is given when $\mathbf{P}, \mathbf{V}, t$ are given, by the identical relation satisfied by the integrals Were ξ a mere parameter, we should have 6 constants plus the value of ξ at our disposal, and the fit could be made anywhere Actually it can only be made at $\xi = \infty$, $|\mathbf{V}| = c$ The reader will find it of great interest to supply the details of this investigation for himself

319. The reader may be puzzled by this situation This is because he has been brought up to believe that differential equations are the fundamental relations governing the universe, and so whenever he sees a differential equation he expects a consequent train of events The situation we are analysing is a totally different state of affairs We are considering primarily not a system of differential equations but a system of trajectories described by moving particles We began our statistical analysis with a statistical set of trajectories, satisfying the cosmological principle We further imposed the condition that no particle was to be created or destroyed From this stage on the complete family of trajectories was determinate It could presumably have been at once handled and described by a sufficiently powerful calculus involving only functional equations Amongst these functional equations is that expressing the permanence of each material particle The poverty of the resources of mathematical analysis led us to express this permanence-condition in differential form, a form we called Boltzmann's equation, and insertion in this of the general functional form of the acceleration-relation led to an unambiguous differential equation The ambiguity of sign only appeared when we integrated this equation in a form involving a square root But no relevant consideration demanded that we should always consider simultaneously both signs of this square root Actually we have to consider sometimes the one and sometimes the other, depending on the trajectory by which the particle arrived at the velocity-event in question A sufficiently powerful calculus avoiding the successive processes of differentiation and integration would have led immediately to the description of the complete trajectories in the forms of the parallel columns of formulae (8)–(16) This system of trajec-

tories is what is the real subject of our discussion, not the incidental differential equations first derived by a differentiation and then in turn painfully integrated in terms of 6 arbitrary constants The sixfold infinity of trajectories were in effect *given* as soon as we had formulated them as satisfying the condition of particle-permanence and the cosmological principle We are discussing phenomena, not differential equations The reader must not allow himself to be humbugged by mathematics Differential equations, like the world in Wordsworth's sonnet, are 'too much with us' The pre-eminent role assigned to differential equations (field equations) in the general theory of relativity is no doubt one of the reasons for the limitations of that theory We have encountered a set of phenomena —trajectories—not describable by determinate differential equations These equations waved a flag to indicate that the situation was difficult for them, and they met the difficulty by offering us a choice of sign in a term, one to be taken in one set of circumstances and another in another

The upshot of this discussion is that the family of trajectories we are examining as a world system is completely determinate, but we cannot embody this determinacy in a single function expressing the acceleration-law in O's experience as a one-valued function of the variables $t, \mathbf{P}, \mathbf{V}$ 'Autres temps, autres mœurs ' The human analyst attempts to describe the trajectory-determinacy in terms of a one-valued acceleration as a function of $t, \mathbf{P}, \mathbf{V}$, but the situation resists description in these terms

Physical interpretation of the second branch

320. As t increases above t_l, on any given trajectory, $|\mathbf{V}|$ (which we now call $|\mathbf{V}'|$) decreases and ξ decreases Clearly as ξ decreases, X' increases and the variable η decreases As $\xi \to 1$, $X' \to \infty$ and $\eta \to 0$ Thus by (12) $t' \to \infty$, and by (11) $|\mathbf{P}'| \to \infty$ But, also, as $\xi \to 1$, $\zeta \to 0$, and by (10) $\mathbf{V}' \to \mathbf{V}'_0$ Thus the physical meaning of the integration constant $\mathbf{V}'_0$ is that it is the limiting velocity, as $t \to \infty$, along the trajectory which began, at $t = 0$, with velocity $\mathbf{V}_0$ Thus every particle acquires ultimately a constant limiting velocity which is a function of its original integration constants $\mathbf{V}_0$ and $\imath$ only, and it ultimately comes to rest relative to that fundamental particle (which we may call P'_0) which moves according to $\mathbf{P}'_0 = \mathbf{V}'_0 t$

In the hodograph diagram (Fig 13) the representative point M

moves from M_0 to M_l in time t_l (in O's experience), then retraces its path and ultimately, as $t \to \infty$, assumes a limiting position M_0'.

Thus after reaching the speed of light, in the second stage of its career, every particle decelerates (at least at first) and ultimately accompanies, in velocity, the nucleus of one of the original sub-systems, but *a different sub-system from that to which it originally belonged*

321. Whether it actually falls into or reaches the nucleus itself, $\mathbf{P}_0'$ $(= \mathbf{V}_0' t)$, we are not in a position to say until the form of $\psi(\xi)$ is more precisely prescribed It would only fall into the nucleus if $\lim_{t'\to\infty}(\mathbf{P}-\mathbf{V}_0' t') = 0$, i e if

$$\lim_{\xi\to 1} X'^{\frac{1}{2}} \sinh \eta = 0,$$

i e if
$$\lim_{\xi\to 1}\left[\frac{1}{2C}\int\limits_1^{\xi}\frac{\xi-1}{\xi^{\frac{1}{2}}}\psi(\xi)\,d\xi\right]\exp\left[\frac{1}{2C}\int\limits_{\xi}^{\infty}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right] = 0$$

Of the factors, the first tends to zero, the second to infinity. The behaviour of the product depends on the nature of the singularity in $\psi(\xi)$ at $\xi = 1$ Obviously a fairly strong singularity in $\psi(\xi)$, yet weak enough to assure the convergence of η to zero, would give a limit ∞ for the product, but a sufficiently mild singularity would permit a non-zero limit, and an exceedingly mild singularity a zero limit It follows that though the particles with a given $\mathbf{V}_0$ and $\mathbf{i}$ come to rest relative to the nucleus of the associated $\mathbf{V}_0'$-sub-system, they may, and in general will, fill a volume of space surrounding the $\mathbf{V}_0'$-nucleus.

322. The complete career of every particle is now patent. If it is the nucleus of a sub-system $\mathbf{V}_0$, it moves for ever with the velocity $\mathbf{V}_0$. If it is not the nucleus of a sub-system, but starts with the velocity $\mathbf{V}_0$ at $t = 0$ (thus 'belonging to' the $\mathbf{V}_0$-sub-system) it accelerates† up to the speed of light, then decelerates and joins the sub-system $\mathbf{V}_0'$ associated with $\mathbf{V}_0$ and $\mathbf{i}$, as given by (21) The value of $\mathbf{V}_0'$ is different for every different $\mathbf{i}$, for given $\mathbf{V}_0$ Thus the non-nuclear members of the $\mathbf{V}_0$-sub-system ultimately join different sub-systems, depending on the values of their integration-constant $\mathbf{i}$ The original sub-system $\mathbf{V}_0$ loses an infinity of members—we may say that it ultimately loses every non-nuclear member—to a definite set of other sub-systems

† Possibly decelerating a little at first, as is evident from the hodograph

For example, our own sub-system, defined by $\mathbf{V}_0 = 0$, loses its members to the sub-systems $\mathbf{V}_0'$ given, by (24), by

$$|\mathbf{V}_0'| = c \tanh 2\eta_l \tag{26}$$

323. But the $\mathbf{V}_0$-sub-system does not go uncompensated for these losses It will, to certain particles of other sub-systems, play the role of a $\mathbf{V}_0'$, and receive from these other sub-systems wanderers which come ultimately to rest relative to itself Putting $\mathbf{V}_0' = 0$ in (24), we find that particles arrive, as $t \to \infty$, at rest relative to our own sub-system, from the sub-systems $\mathbf{V}_0$ given by

$$|\mathbf{V}_0| = c \tanh 2\eta_l \tag{27}$$

Comparison with (26) gives at once a particular case of certain reciprocal relations we shall establish shortly For the '$\mathbf{V}_0 = 0$' sub-system ultimately loses members to those precise sub-systems from which it ultimately gains members

324. The foregoing translation of our mathematics into words must not be allowed to mislead The actual state of affairs in the universe will be complicated by the local accelerations, tidal couples, and rotations of the galaxies or sub-systems arising from the imperfect way in which they satisfy the cosmological principle, due in turn, most probably, to that discreteness in three dimensions which is not fully described by our statistical, average-taking methods What we have been describing are our average expectations Every member of a sub-system is accelerated away from the nucleus of the sub-system to which it belongs, and towards the apparent centre in the frame in which it is momentarily at rest, passes through the speed of light, and ultimately joins (in velocity) some other sub-system Its acceleration then tends to zero, and it may be said ultimately to succeed in its aim of falling into the apparent centre in the frame in which it ultimately comes to rest

Associated sub-systems

325. The sub-system $\mathbf{V}_0'$ where $\mathbf{V}_0'$ is given by (21) may be said to be associated with the sub-system $\mathbf{V}_0$ in respect to the vector $\mathbf{i}$ It is easy to calculate the *present* positions of the nuclei of the sub-systems associated with ourselves ($\mathbf{V}_0 = 0$), whose members ultimately come to rest relative to ourselves It follows from (27) that their present distance from us is

$$ct \tanh 2\eta_l, \tag{28}$$

where t is our present epoch. But a less short cut to the result gives more insight. For trajectories along which the corresponding particles come ultimately to rest relative to ourselves, $\mathbf{V}_0'$ must be zero. Such a trajectory 'belongs to' a sub-system $\mathbf{V}_0$ given by (22). But since $\mathbf{V}_0' = 0$, we have $\epsilon' = 0$, and hence by (22)

$$\mathbf{V}_0 = ct\mathbf{i}\tanh 2\eta_l. \tag{29}$$

Hence the nucleus of the sub-system to which the ultimately arriving particle originally belonged is now at a position $ct\tanh 2\eta_l\,\mathbf{i}$ from ourselves. Such particles ultimately arrive from the direction $\pm\mathbf{i}$, for by (10) their velocity $\mathbf{V}'$ is parallel to $\pm\mathbf{i}'$, i.e. to $\pm\mathbf{i}$, the sign depending on the relative magnitudes, for $\xi \sim 1$, of η and ζ. From (29) it follows that the integration parameter ϵ for such particles is $2\eta_l$. For

$$\sinh\epsilon = \frac{(\mathbf{V}_0\cdot\mathbf{i})/ct}{(1-V_0^2/c^2t^2)^{\frac{1}{2}}} = \sinh 2\eta_l.$$

It is worth noting that the same result follows, as it should, from (21). Putting $\mathbf{V}_0' = 0$ in (21) we have

$$\mathbf{V}_0 = ct\mathbf{i}(1-V_0^2/c^2t^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)}. \tag{29a}$$

Hence $\mathbf{V}_0$ is parallel to $\mathbf{i}$. Multiplying scalarly by $\mathbf{i}$ we have at once

$$\sinh\epsilon = \frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)},$$

whence

$$\epsilon = 2\eta_l, \tag{29b}$$

and (29) follows.

326. The fundamental character of the world-constant η_l is evident. The results we have just obtained should be contrasted with our results on cosmic rays in the preceding chapter. Galaxies from which we are *now* receiving swift particles are at a present distance $ct\tanh\eta_l$, galaxies from which we shall *ultimately* receive particles coming asymptotically to relative rest are at a present distance $ct\tanh 2\eta_l$. Both these distances are less than the radius of the expanding light-sphere $\mathbf{P} = ct$, and the first distance is less than the second.

Reciprocity theorems

327. We shall now establish some remarkable reciprocal relations.

Consider first the trajectory defined, in O's experience, by any $\mathbf{V}_0$, any $\mathbf{i}$, and any X_l. It possesses an associated constant ϵ, a known

function of $\mathbf{V}_0$ and $\mathbf{i}$ Its second branch possesses constants $\mathbf{V}_0'$, $\mathbf{i}$, and X_l, where $\mathbf{V}_0'$ is given in terms of $\mathbf{V}_0$ and $\mathbf{i}$ by any of (20), (21), (22)

Consider next a trajectory of any other X_l, whose '$\mathbf{V}_0$' (velocity at $t = 0$) has the value $\mathbf{V}_0'$ just calculated, and whose '$\mathbf{i}$' is $-\mathbf{i}$, where $\mathbf{i}$ is the value for the trajectory just considered The 'ϵ' of the new trajectory is accordingly $-\epsilon'$, where ϵ' is the constant associated with the second branch of the first trajectory The new trajectory on its second branch will possess constants which we will call $\mathbf{V}_0''$ and $\mathbf{i}''$, with an associated ϵ'' Then $\mathbf{i}''$ and ϵ'' are given, by (17) and (18), by

$$\mathbf{i}'' = -\mathbf{i}, \tag{30}$$

$$-\epsilon'-\epsilon'' = 2\eta_l \tag{31}$$

Hence
$$\epsilon'' = -2\eta_l-\epsilon'$$
$$= -\epsilon, \tag{32}$$
by (18)

Hence the ultimate velocity along the second trajectory, namely $\mathbf{V}_0''$, is given, by (21), by

$$\mathbf{V}_0'' = \mathbf{V}_0'-c(-\mathbf{i})(1-V_0'^2/c^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(-\epsilon'-2\eta_l)}$$
$$= \mathbf{V}_0'+c\mathbf{i}(1-V_0'^2/c^2)\frac{\sinh 2\eta_l}{\cosh(\epsilon'+2\eta_l)}$$
$$= \mathbf{V}_0, \tag{33}$$
by (22)

Thus, whilst the trajectory $(V_0, \mathbf{i}, \epsilon)$ has for its second branch $(\mathbf{V}_0', \mathbf{i}, \epsilon')$, the trajectory $(\mathbf{V}_0', -\mathbf{i}, -\epsilon')$ has for its second branch $(\mathbf{V}_0, -\mathbf{i}, -\epsilon)$, independent of the chosen values of the X_l's We may write this symbolically thus

$$\left.\begin{aligned} (\mathbf{V}_0, \mathbf{i}, \epsilon) &\to (\mathbf{V}_0', \mathbf{i}, \epsilon'), \\ (\mathbf{V}_0', -\mathbf{i}, -\epsilon') &\to (\mathbf{V}_0, -\mathbf{i}, -\epsilon) \end{aligned}\right\} \tag{34}$$

It follows that for every particle which the $\mathbf{V}_0'$-sub-system ultimately gains from the $\mathbf{V}_0$-sub-system, the $\mathbf{V}_0$-sub-system ultimately gains a particle from the $\mathbf{V}_0'$-sub-system

328 Consider this pair of trajectories more closely For given X_l, the $(\mathbf{V}_0, \mathbf{i})$ particle attains the speed of light at a definite epoch and in a definite position Let $X_l^{(1)}$ be the value of X_l for this trajectory Let $X_l^{(2)}$ be the chosen value of X_l for the trajectory $(\mathbf{V}_0', -\mathbf{i})$ Along this second trajectory the corresponding particle also attains the speed

of light at some definite epoch and in some definite position. We shall prove that if $X_l^{(1)} = X_l^{(2)}$, the two trajectories intersect at a common epoch, and that this epoch, and the associated point of intersection, are the epoch and position at which the two particles attain the speed of light. They are then moving in different directions, and may be said to collide.

329. For the time to light velocity on the $(X_l^{(1)}, \mathbf{V}_0, \mathbf{i})$ trajectory is, by (12),

$$t_l^{(1)} = \frac{X_l^{(1)\frac{1}{2}}}{(1-V_0^2/c^2)^{\frac{1}{2}}}\,\frac{\cosh(\epsilon-\eta_l)}{\cosh\epsilon}. \tag{35}$$

The time to light velocity on the $(X_l^{(2)}, \mathbf{V}_0', -\mathbf{i}')$ trajectory is by the same formula

$$t_l^{(2)} = \frac{X_l^{(2)\frac{1}{2}}}{(1-V_0'^2/c^2)^{\frac{1}{2}}}\,\frac{\cosh(-\epsilon'-\eta_l)}{\cosh(-\epsilon')}. \tag{36}$$

But from (18),

$$\cosh(\epsilon-\eta_l) = \cosh(\epsilon'+\eta_l) = \cosh(-\epsilon'-\eta_l).$$

Using this and (23), we see that the two times $t_l^{(1)}$ and $t_l^{(2)}$ are equal provided

$$X_l^{(1)} = X_l^{(2)}. \tag{37}$$

330. Again, the position $\mathbf{P}_l^{(1)}$ at which the particle on the $(X_l^{(1)}, \mathbf{V}_0, \mathbf{i})$ trajectory attains the speed of light and the corresponding epoch $t_l^{(1)}$ satisfy the relation

$$\frac{\mathbf{P}_l^{(1)}}{t_l^{(1)}} = \mathbf{V}_0 - c(1-V_0^2/c^2)^{\frac{1}{2}}\,\frac{\sinh\eta_l}{\cosh(\epsilon-\eta_l)}\,\mathbf{i} \tag{38}$$

by the first of formulae (11) and (12). Consequently along the $(X_l^{(2)}, \mathbf{V}_0', -\mathbf{i})$ trajectory the position $\mathbf{P}_l^{(2)}$ and epoch $t_l^{(2)}$ at which the speed c is attained are connected by

$$\frac{\mathbf{P}_l^{(2)}}{t_l^{(2)}} = \mathbf{V}_0' - c(1-V_0'^2/c^2)^{\frac{1}{2}}\,\frac{\sinh\eta_l}{\cosh(-\epsilon'-\eta_l)}(-\mathbf{i}). \tag{39}$$

By (20), these are equal. But we have just proved that $t_l^{(1)} = t_l^{(2)}$. Hence $\mathbf{P}_l^{(1)} = \mathbf{P}_l^{(2)}$, and the positions coincide.

331. We notice further that $\mathbf{P}_l/t_l$ is independent of the value of X_l, and accordingly all trajectories with given $\mathbf{V}_0, \mathbf{i}$ attain the velocity of light at points whose locus is a straight line, namely a vector in a fixed direction from O given by (38), and the distances along this straight line, for varying X_l, are proportional to t_l.

332. We have throughout been considering the forms of trajectories calculated as if unimpeded by collisions We have agreed to regard collisions as secondary effects outside our ideal scheme of motions, modifying its realization in nature We are not to suppose that given an $(X_l, \mathbf{V}_0, \mathbf{i})$ trajectory there is really existing the precise trajectory $(X_l, \mathbf{V}'_0, -\mathbf{i})$, for the probability that the exact conditions $X_l^{(2)} = X_l^{(1)}$, $\mathbf{i}^{(2)} = -\mathbf{i}^{(1)}\ (= -\mathbf{i})$, $\mathbf{V}_0^{(2)} = \mathbf{V}'_0$, $\mathbf{V}_0^{(1)} = \mathbf{V}_0$ are satisfied by some pair of trajectories is zero It is an unavoidable limitation of the statistical method of analysis that it treats as existents relations that will not be exactly satisfied Thus our analysis is not to be taken to mean that this careful timing of motions which theoretically results in collisions is exactly realized

333. A still further result of interest may be proved We shall show that $\mathbf{V}'_0$ lies inside the angle formed by $\mathbf{P}_l$ and $\mathbf{V}_l$ This will also imply that $\mathbf{V}_0$ lies inside the angle formed by $\mathbf{P}_l^{(2)}$ $(= \mathbf{P}_l)$ and $\mathbf{V}'_l$

For we have

$$\mathbf{V}_l = \mathbf{V}_0 - c\mathbf{i}(1-V_0^2/c^2)^{\frac{1}{2}}(\cosh\epsilon+\sinh\epsilon), \tag{40}$$

$$\frac{\mathbf{P}_l}{t_l} = \mathbf{V}_0 - c\mathbf{i}(1-V_0^2/c^2)^{\frac{1}{2}}\frac{\sinh\eta_l}{\cosh(\epsilon-\eta_l)}, \tag{41}$$

$$\mathbf{V}'_0 = \mathbf{V}_0 - c\mathbf{i}(1-V_0^2/c^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)} \tag{42}$$

The desired result will follow if we can show that

$$\frac{\sinh\eta_l}{\cosh(\epsilon-\eta_l)} < \frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)} < \cosh\epsilon+\sinh\epsilon \tag{43}$$

These inequalities are readily established

Hodograph and position diagrams

334. We are now ready to exhibit the beautiful geometrical relationships of our trajectories in a position-hodograph diagram

In Fig 19, O represents the observer, and the figure represents the state of affairs in O's experience OM_0 represents the vector $\mathbf{V}_0$, OM'_0 the vector $\mathbf{V}'_0$ M'_0M_0 is the sense of the unit-vector $\mathbf{i}$ M_l, M'_l are points on the hodograph $M_0M'_0$ such that $OM_l = OM'_l = c$ $\mathbf{P}_0$ is the point $\mathbf{V}_0 t_l$, $\mathbf{P}'_0$ the point $\mathbf{V}'_0 t_l$, and the line joining P_0, P'_0 is parallel to $-\mathbf{i}$ The position P_l at which the particle following the $(X_l, \mathbf{V}_0, \mathbf{i})$ trajectory attains the speed of light lies on $P_0P'_0$, and this is the same point P'_l (previously called $P_l^{(2)}$) at which the particle

following the associated trajectory $(X_l, \mathbf{V}_0', -\mathbf{i})$ attains the speed of light These two trajectories are shown in the diagram

Along the trajectory OP_l, which is tangent to the vector $\mathbf{V}_0$ at O, the representative point M in the hodograph moves from M_0 to M_l, and the tangent at P_l is parallel to OM_l After P reaches P_l, it follows a trajectory which becomes ultimately parallel to $\mathbf{V}_0'$, and the representative hodograph point M retraces its path from M_l along $M_l M_0'$, coming to rest at M_0' as $t \to \infty$

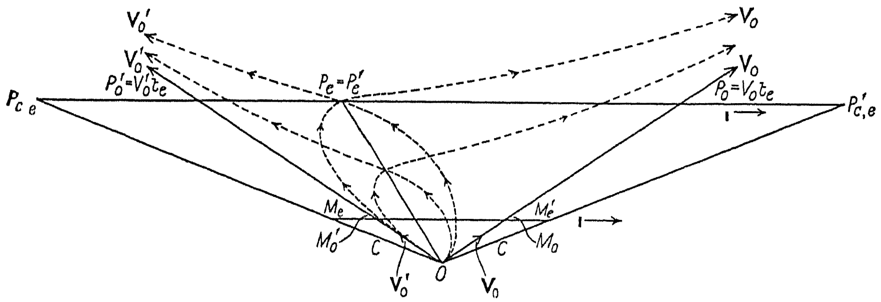


FIG 19 Complementary trajectories and their passage through the speed of light

Along the trajectory OP_l' (where $P_l' \equiv P_l$), which is tangent to $\mathbf{V}_0'$ at O, the representative point M' in the hodograph moves from M_0' to M_l', and the tangent at P_l' is parallel to OM_l' After P reaches P_l', it follows a trajectory which becomes ultimately parallel to $\mathbf{V}_0$, and the representative hodograph point M' retraces its path from M_l' along $M_l'M_0$, coming to rest at M_0 as $t \to \infty$

The particle following the trajectory OP_l which started with the nucleus of the sub-system following the straight line $\mathbf{P}_0 = \mathbf{V}_0 t$ ultimately accompanies in velocity the nucleus of the sub-system following the straight line $\mathbf{P}_0' = \mathbf{V}_0' t$ Converse relations hold for the trajectory OP_l'

335. A second pair of associated trajectories is shown in the figure, defined by the same values $\mathbf{V}_0$ and $\mathbf{i}$, or $\mathbf{V}_0'$ and $-\mathbf{i}$, but possessing a different value for X_l ($= X_l'$) From what has been proved, these trajectories also intersect on the straight line OP_l, and their particles pass through the point of intersection simultaneously, the epoch of 'collision' bearing to t_l the same ratio as the distance of the point of collision bears to OP_l

336 Every trajectory has a point of inflexion at its point P_l, where the particle-velocity is c, the two branches of the trajectory are the

two solutions of the two differential equations associated with opposite signs for the surd

337. The diagram clearly exhibits the interchange of members of the sub-systems $\mathbf{V}_0$ and $\mathbf{V}_0'$. It must be remembered that $\mathbf{V}_0$ and $\mathbf{V}_0'$ are not both arbitrary. The possible values of $\mathbf{V}_0'$, for given $\mathbf{V}_0$, form a two-parameter system of vectors given by (21), the system depending on the world-constant η_l. Thus for given $\mathbf{V}_0$, the extremities of the vectors $\mathbf{V}_0'$ lie on a surface, which is a surface of revolution around $\mathbf{V}_0$ as axis. The shape of the surface, in O's experience, depends on η_l and on $|\mathbf{V}_0|/c$. For $\mathbf{V}_0 = 0$, it reduces to the sphere

$$\mathbf{V}_0' = -c\mathbf{i}\tanh 2\eta_l$$

338. The diagram displays OP_0' (the direction of $\mathbf{V}_0'$) as lying between OP_l and OM_l, and similarly OP_0 as lying between OP_l' ($= OP_l$) and OM_l', in accordance with what we have proved. For fixed $\mathbf{V}_0$ and varying $\mathbf{i}$, the points M_l lie at the ends of the chords $\pm\mathbf{i}$ (drawn through the extremity M_0 of the vector $\mathbf{V}_0$) of the sphere of centre O and radius c. The positions $P_{c,l}$, $P_{c,l}'$ of the apparent centres of the system to the two particles passing through P_l ($= P_l'$), in the frames in which they are momentarily at rest, are the intersections of the chord in direction $\pm\mathbf{i}$, drawn through P_0 with the sphere of centre O and radius ct

339 For given $\mathbf{V}_0, \mathbf{i}$, the time t_l to light-speed is simply proportional to $X_l^{\frac{1}{2}}$. This can take any value between 0 and ∞. Thus particles are arriving in any given vicinity with speeds approximating to that of light at all times

340 The whole of the relations expressed by the above diagram are invariant under transformation from the particle-observer O to any other fundamental particle-observer O', for all are equivalent. Each observer O' can make up his own diagram for the pair of particles following associated trajectories just considered. No two such diagrams coincide, but all have the same general form and the same properties. A particle which is passing through the speed of light for any one particle-observer O is also passing through the speed of light for any other fundamental particle-observer O'

341. The detailed shapes of the trajectories depend on the precise form of $\psi(\xi)$, which we have left arbitrary subject to certain types of

behaviour at $\xi \to 1$ and $\xi \to \infty$ It is remarkable that so much can be established without further information as to $\psi(\xi)$ It means that the properties we have established are common properties of a continuum of universes, or rather more than a continuum of universes, for the cardinal of the class of continuous functions exceeds the cardinal of the continuum

342 Whilst P travels from O to P_l, the apparent centre (to P) of the whole system travels from O to $P_{c,l}$ by a path along which the vector PP_c steadily increases in length At $P = P_l$, it has a maximum Beyond P_l, it steadily decreases, at least at first It is easily found that on the second branch,

$$\frac{d}{dt}|\mathbf{P}'-\mathbf{V}'t| = -\frac{t'Y'}{X}|\mathbf{P}'-\mathbf{V}'t'|\left[-1+\frac{C}{(\xi-1)^{\frac{3}{2}}\psi(\xi)}\right],$$

so that $|\mathbf{P}'-\mathbf{V}'t'|$, the distance of P from the apparent centre, decreases so long as

$$(\xi-1)^{\frac{3}{2}}\psi(\xi) < C$$

How far this holds, and whether it always holds, depends on the nature of the singularity in $\psi(\xi)$ at $\xi = 1$ This singularity might be further investigated by discussing in detail the coexistence of the relations

$$\lim_{t\to\infty} X' = X'_l \lim_{\xi\to 1} \exp\left[+\frac{1}{C}\int_{\xi}^{\infty}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right], \quad \xi = \frac{Z'^2}{X'Y'}, \quad Y' \to 1-\frac{V_0'^2}{c^2}$$

for the various possible forms of $\psi(\xi)$ near $\xi = 1$ which are compatible with the divergence at $\xi = 1$ of $\int_{\xi}^{\infty}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi$ and the convergence at $\xi = 1$ of $\int_{\xi}^{\infty}\frac{\xi-1}{\xi^{\frac{1}{2}}}\psi(\xi)\,d\xi$. But I do not propose to undertake such investigations here My conjecture is that $\mathbf{P}'-\mathbf{V}'t' \sim \mathbf{P}'-\mathbf{V}_0't' \to$ constant as $\xi \to 1$, $t' \to \infty$, but the grounds for this will appear later If we could assume this, it would follow from (11) and (13) that $X'^{\frac{1}{2}}\eta$ and $X'^{\frac{1}{2}}\zeta$ tend to finite limits as $\xi \to 1$, $\zeta \to 0$, $\eta \to 0$, and this would fix the form of the singularity in $\psi(\xi)$ as $\xi \to 1$ as given by

$$\exp\left[+\frac{1}{2C}\int_{\xi}^{\infty}(\xi-1)^{\frac{1}{2}}\psi(\xi)\,d\xi\right] \sim \frac{\text{const}}{(\xi-1)^{\frac{1}{2}}}, \qquad (\xi \sim 1),$$

or

$$\psi(\xi) \sim \frac{C}{(\xi-1)^{\frac{3}{2}}}, \qquad (\xi \sim 1)$$

This satisfies all our conditions, but in the time at my disposal I have not seen my way to establish it *a priori* If this result holds good, we may say that the complete motion of P is one of pursuit of the apparent centre (to it) of the universe, following projection away from it with velocity $\mathbf{V}_0$, and of the ultimate practically successful overtaking of the apparent centre as $t \to \infty$ The whole history of the totality of trajectories then appears in a rational and easily intelligible light.

343 It should be remembered that it is *only in the experience of O* that the gravitational pull of the system of the universe always changes sign as P's velocity passes through c In P's experience the acceleration of P reckoned relative to the moving apparent centre P_c (not the fundamental particle with which P_c happens instantaneously to coincide) is always *towards* P_c But the exact mathematical embodiment of P's experiences of the system would require the use of formulae of transformation from an O to an accelerated P, which we have not here investigated We have not found it necessary to use any formulae of transformation save from an O to an equivalent O' Formulae of transformation from coordinates used by O to coordinates used by P (from P's own measures) are of an altogether different order of profundity

Astronomical identification Cosmic clouds

344 We now inquire whether there is any evidence in nature of the remarkable result predicted by our analysis, namely, the ultimate arrival in the vicinity of each nebular nucleus of particles coming to rest relative to that nucleus Is there any evidence of particles from outside coming to rest relative to ourselves as time (in our experience) advances? Without such evidence the discussions of the present chapter, though of fascinating mathematical interest, would be largely academic

345. The phrase '$t' \to \infty$' must not be misunderstood We have introduced no absolute reckoning of any 'large' time—there is no physical constant, of the dimensions of a time, in our work in comparison with which t must be large Whatever happens in the limit, as $t' \to \infty$, must be happening, approximately, for some particles for any given t' This follows at once from our formulae If we consider our own vicinity, which is given by ξ near 1, or η small, and ask for particles coming to rest relative to our own galaxy, as $t \to \infty$, then at

our present epoch t, they will already have very small velocities provided $(\eta-\zeta)$ is sufficiently small, as follows from (10) on putting $\mathbf{V}_0' = 0$ and determining $\mathbf{V}'$ By (12) we can take t' equal to our present epoch by appropriate choice of X_l, and then from (11), putting $\mathbf{V}_0' = 0$, $\mathbf{P}'$ will also be 'small' if $\psi(\xi)$ has the right kind of singularity Hence by any given time t a number of particles bound for $\mathbf{V}_0' = 0$, i e for rest relative to us, will already have been reduced to small velocities Such particles should have congregated in our own vicinity to some extent at any given epoch, and their presence should be apparent amongst the surroundings or contents of our own galaxy It should already have gained some members from external galaxies, and these gained members should be sharing the motion of our own galaxy relative to neighbouring galaxies

346 A survey of our own galaxy immediately discloses objects at rest relative to our galaxy as a whole, namely the interstellar cloud or clouds of calcium and sodium vapour, and the cosmic dust-clouds

It was discovered in 1905 by Hartmann that the absorption lines due to ionized calcium in the spectrum of the star δ Orionis, which is a spectroscopic binary, did not share the 'motion' of the other lines in the spectrum, but remained 'stationary' He attributed the origin of the stationary lines to a cloud of calcium vapour surrounding the double star and at rest relative to its centre of mass Since then, 'stationary' lines of calcium and sodium have been discovered in the spectra of many other binary stars, but in all cases the stars are of type $B\,3$ or earlier There have also been observed lines of the same character—sharp and narrow—in non-binary stars, in such case the radial velocity indicated by the calcium lines does not necessarily coincide with that indicated by the other lines in the spectrum It was shown by J S Plaskett in 1924 that the source of the 'stationary' lines has a mean radial velocity agreeing with the reflex of the radial component of the sun's motion amongst the stars It was further shown by J S Plaskett in 1929 that the motions displayed by the radial velocities for the stationary lines indicated a rotation of the source about a centre closely agreeing with the direction of the centre of galactic rotation discovered by Oort in 1927, and that the mean velocities for separate groups of 'stationary' lines agreed well with the hypothesis of galactic rotation The still more significant result was discovered by him that the 'rotational term' (in Oort's

notation) for the source of the 'stationary' lines is very closely one-half that for the corresponding stars, i e that the mean source of the lines in question is one-half the distance of the stars

These various results can be reconciled by the assumption (due to Plaskett)† of the existence of a widely diffused cloud of 'vapour' containing ionized calcium and sodium atoms, more or less regularly distributed and participating in the galactic rotation That the 'stationary' lines are only observed in the hotter stars is explained by the fact that in the cooler stars any absorption due to the calcium cosmic cloud would be masked by the calcium in the stellar atmospheres, in still hotter stars, on the other hand, calcium in their atmospheres will be doubly ionized and the atmospheres are incapable of showing the H and K lines, in accordance with the theory of high-temperature ionization The physical properties of such a 'gaseous substratum' in the galaxy were studied in detail by Struve and Gerasimovič in 1929 Making allowances for the ionization of the calcium and sodium and for the probable partial abundance of calcium in relation to other constituents, they concluded that its density was of the order of 10^{-26} gramme cm $^{-3}$, a little higher than the range 10^{-27}–10^{-30} of the present mean local density of matter in space

347. Such a gaseous interstellar cloud of ionized and other atoms thus shares the mean motion of the stars embedded in it It is 'at rest' relative to the stars of our galaxy as a whole, in the two senses that it yields the same value for the solar motion and the same value for the galactic rotation

348. It is far from certain that this interstellar cloud has not originated from the stars of our galaxy themselves In early papers‡ on the solar chromosphere I pointed out that the selective radiation pressure on ionized calcium atoms in the atmospheres of the hotter stars would make it difficult for them to form a screen of such atoms of such a thickness that the outermost atoms would be in mechanical equilibrium, and that in that case the atoms would be expelled under radiation pressure, thereby tending to populate interstellar space I later§ indicated other mechanisms by which atoms might be expelled from even cooler stars like the sun Processes of this kind have been

† Following a theoretical investigation by Eddington, which has, however, been sharply criticized by Jeans

‡ *M N , R A S* , **84,** 361, 1924

§ Ibid , **86,** 459, 1926

considered in great detail in more recent years by Beals, Chandrasekhar, and others, and it is quite possible that the totality of such processes in past ages has furnished the constituents of the interstellar cloud It seems, indeed, reasonably certain that the interstellar cloud has been and is being recruited by such processes The atoms emitted would possess in the mean the mean motion of the stars emitting them, as observed

349. Nevertheless the estimated density of the interstellar cloud seems somewhat high for this process to be the sole agency. Here I would only point to the possibility that the interstellar cloud includes arrivals from other galaxies The ultimately arriving particles of our general theory would share the motion of the centre of the galaxy, they would be free from any K-effect, unlike the stars or other proper members of our sub-system, and this is precisely what J S Plaskett found for the calcium atoms—zero K-effect Further, having come into our vicinity, they would be subject to any local irregularities in the mean gravitational field of the cosmos present in our neighbourhood due to chance fluctuations from the idealized distribution of galaxies, and if it is their fluctuations which are responsible for the rotation and flattening of our own system (if flattened it be), then the arriving atoms would tend to participate in time in the galactic rotation

350. But I do not press this possible explanation For there is a totally separate body of evidence as to the presence in our galaxy of a stratum of material, an 'obscuring dust-cloud' extending over and parallel to the galactic equator The researches of Trumpler, Bok, van de Kamp, Stebbins, Hubble, and others have shown that there is a discrepancy between the observed colours of the stars and their observed spectral-types which is a function of galactic latitude and indicates a reddening of stars in low galactic latitudes, that this extends to the globular clusters of stars which are supposed to outline our own galaxy, that counts of extra-galactic objects indicate a general obscuration in the galactic plane, by which there is a systematic decrease in the number of objects of a given apparent brightness from galactic pole to equator, that this obscuration is practically complete in the galactic plane itself, for extra-galactic nebulae are not observed within a few degrees on either side of this plane, and that these various results may be harmonized by the assumption of a relatively thin

stratum of obscuring material, contributing an optical thickness of perhaps 0 8 photographic magnitude (van de Kamp) or 0 5 magnitude (Hubble) The sun is situated inside the layer, and roughly half-way between its boundaries, the layer seems to be symmetrical about the galactic equator The obscuration of objects in higher galactic latitudes is expressed fairly accurately by a cosecant formula, indicating a plane parallel-sided slab of dust particles

351. As van de Kamp has pointed out, and as must have occurred to many, the existence of an obscuring belt is not peculiar to our own system alone It has been known for long that many spiral nebulae, seen edgewise, show well-marked equatorial bands of opaque matter Indeed, Curtis in 1918 pointed to the occurrence of dark lanes in spiral nebulae as a relatively common phenomenon, and considered it as evidence for a similar band of occulting matter at the periphery of our own system To van de Kamp it suggested rather a more or less continuous equatorial layer, and the suggestion would appear to be generally acceptable Quantitative comparisons of the size of the obscuring layer in other systems with that in our own confirm the suggestion Estimates of the linear thickness of the layer range from 1,600 parsecs to 400 parsecs, the lower value being favoured †

352 Much work is now being done on the subject, and the above summary is very incomplete It is not our object to treat in detail of the whole existing observational material But enough has been said to justify raising the question what is the origin of this obscuring material, present in our own system and in distant galaxies ?

353. I suggest that our mathematical analysis furnishes an answer I suggest that the cosmic dust-clouds present in and around the galaxies are the products of the arrivals, in the vicinity of nebular nuclei, of particles from other galaxies I suggest that we identify the ultimately arriving particles, which as time advances come to rest relative to, and in the vicinity of, each galaxy as predicted by the general theory, with the constituents of the obscuring clouds, not excluding also a possible identification with the constituents of the interstellar calcium and sodium clouds, whose relation to the dust-clouds is not yet understood ‡

† Trumpler, *L O B*, **14**, 154, 1930 Bok, *Harvard Observatory Circular*, No 371, 1931 van de Kamp, *Astronom Journ*, **42**, 1932 Stebbins, *Proc Nat Acad Sci*, **19**, 222, 1933 Hubble, *Astrophys Journ*, **79**, 8, 1934

‡ See the paper of E G Williams, *Astrophys Journ*, **19**, 280, 1934

N G C 4565, *Coma Berenices*, Spiral Nebula on edge, H V 24, exposure 5 hours, March 6 and 7, 1910 60 inch Reflector

(By the courtesy of the Director of the Mount Wilson Observatory)

354 There is no direct evidence of the motions of the 'cosmic' clouds or obscuring clouds associated with the different galaxies But their mere existence in association each with its own galaxy shows that the 'proper motion' of each obscuring cloud coincides with that of the associated nucleus Thus each cosmic cloud is at rest relative to the corresponding nucleus, and hence its separate particles share the motion of the corresponding nucleus The obscuring cloud may, like the calcium cloud, share the rotational motion of the associated galaxy, or it may not This would be, in my view, a secondary or induced phenomenon The main point is that observational astronomy, in three different directions—(1) the interstellar calcium cloud, (2) the galactic layer of obscuring material giving rise to colour excesses in stars and reduction even to ultimate complete obscuration of the brightness of objects towards the galactic equator, and (3) the obvious equatorial belts of obscuring matter seen in not too distant external galaxies—indicates the presence of particles at relative rest with respect to the associated systems

355. It has been often conjectured that the stars have been formed as condensations in widely extended primeval distributions of matter We have regarded them as possibly formed by agglomerations resulting from successive collisions of primary 'particles'—a not essentially different origin In either case the system of the stars forming a galaxy would appear to be older than the obscuring clouds themselves, for if they were coeval, the clouds should have condensed or concentrated by now equally into stars I do not remember to have seen this point mentioned before If then the cosmic dust-clouds are relatively younger, they must have been formed after the galaxy had already evolved into a set of stars Our explanation of them as the result of arrivals from other galaxies is in harmony with this circumstance Our explanation predicts that they are still, and ever will be, in process of formation and growth, and that the oldest galaxies, namely ourselves and those in our own neighbourhood (reckoned as seen by ourselves), will possess the most developed cosmic clouds Thus our proposed identification of the ultimately arriving particles from distant nebulae with the material of the obscuring clouds simultaneously accounts for their formation, motion, and non-resolution into stars Obscuring dust-clouds are the accumulated and accumulating debris of the disintegration of distant galaxies by the

pull of the whole universe, and this same pull brings them to 'rest' relative to some nuclear particle or other, statistically Each particle is endowed ultimately with some constant limiting velocity, the value depending on the initial circumstances Each ultimately arriving particle has passed through a phase of its trajectory where its velocity was comparable with that of light, and only those particles arrive in this manner which have happened to avoid serious collisions, in particular have not taken part in a cosmic ray phenomenon

356. We did not set out to seek any *ad hoc* explanation of cosmic dust-clouds The conclusion that particles ultimately attain constant limiting velocities was, however, an inevitable consequence of our formal analysis, and unless we could point to such particles in existence near galaxies the complete identification of our theoretical structure with the system of the universe would have been wanting We find that such particles are in fact present, and this confirms our general scheme This particular feature of the general scheme provides in turn an explanation of the origin, continued existence, and probable future of the cosmic dust-clouds which is entirely unforced

357. There are various features to notice about our explanation Our analysis has been throughout statistical We have always tacitly employed volumes large enough to contain a large number of units Thus when we say that, as $t \to \infty$, a particle tends to assume the velocity $\mathbf{V}_0'$ and so comes ultimately to rest relative to the $\mathbf{V}_0'$-nebula, we mean that the volume of space associated with the recession velocity $\mathbf{V}_0'$ will tend to be tenanted with such particles In nature not every possible value of $\mathbf{V}_0'$ will be represented by a nebular nucleus, but only a discrete set of values of $\mathbf{V}_0'$ The ultimately arriving particles, being more numerous than the nuclei themselves, will more nearly represent a continuum of velocity There will therefore be some dispersion of the ultimately arriving particles in velocity, and an assigned $\mathbf{V}_0'$-nebula will collect particles for the range of $\mathbf{V}_0'$ represented by this nebula in relation to its nearest neighbours We may say that the quasi-continuum of ultimate velocities $\mathbf{V}_0'$ will be divided up into discrete, finite domains, each domain corresponding to some existing sub-system or nebula

358. The observational evidence summarized above suggests that the limiting position P' of a particle is such that $\mathbf{P}'-\mathbf{V}_0't'$ or $\mathbf{P}'-\mathbf{V}'t'$

tends to a non-zero constant For the dust-particles which have already arrived would appear to be each at a definite non-zero distance from the galactic centre In this case, as we saw, further information about the function $\psi(\xi)$ can be obtained, namely the nature of its singularity at $\xi = 1$

359 As regards the galactic flattening of the cosmic dust-clouds, it is reasonable to suppose that this has been caused in some way by the same 'forces', to use a dynamical mode of thought, as have produced the characteristic detailed forms of the nebulae Our statistical analysis only enables us to calculate average effects, and it gives spherical symmetry for the arriving particles, that is, it predicts that the average dust-cloud obtained by superposing all the dust-clouds in a given not too small volume would be spherical in form It is, of course, remarkable that dust-clouds should be so strongly concentrated towards the galactic planes That the concentration should be more marked than for the stars in general is, however, a general consequence of our results For whilst the dust-cloud particles come to relative rest, the stars themselves are animated by an outward motion and an outward acceleration, and thus are subject to a tendency to diffuse outwards

360. The density of dust-clouds could be calculated much as we have calculated the intensity of cosmic rays and the density-distribution near a galactic-nucleus We should again use Jacobians to transform from our distribution in space and velocity to the space-distribution of the ultimately arriving particles in terms of the constants of integration, and then transform back to space-coordinates, lastly we should put $V_0' = 0$, and examine the behaviour for $\xi \sim 1$ and ξ decreasing, at any arbitrary epoch t This would be a most fascinating calculation, but I have not yet had time to carry out the details

361. If the obscuring clouds consist of dust-particles, we may identify at least some of our ideal particles as dust-particles On dynamical grounds (the equivalence of gravitational and inertial mass) we should expect particles of all sizes to follow the same forms of trajectories The dust-particles may range from particles of macroscopic size down to molecules, atoms, or even neutrons or other elementary particles Generally speaking, then, it is convenient to regard our ideal particles as typical dust-particles or smaller entities Larger agglomerations,

formed by collisions, may be expected also to arrive from outside at relative rest as $t \to \infty$, but such visitants should be much rarer than dust-particles It is tempting to speculate that some of the objects in the galactic system possessing low absolute velocities relative to the galactic centre have also arrived from the outside, and will accompany us in perpetuity, they are to be distinguished from the other casual visitants in course of passage through the galactic system (also arrivals from outside) which possess non-zero velocities relative to the centre

362 The foregoing explanation has not been previously suggested and I would repeat that it is not an *ad hoc* explanation, devised with the object of explaining the facts The ultimate arrival from other galaxies of particles coming to relative rest is an intrinsic part of the scheme of phenomena exhibited by a statistical system satisfying the cosmological principle, and in this as in so many other ways it is confirmed by observation I am aware that various arguments could be advanced against my identification, and that apparently conflicting facts might be cited But the subject is in a state of rapid development, and a provisional explanation is better than none Rutherford once remarked that no phenomenon was ever explained unless a simple hypothesis was adopted soon after its discovery, as the multiplicity of new facts rapidly discovered would otherwise negative the chance of any explanation whatever receiving attention Accordingly I submit my explanation of cosmic dust-clouds, fully prepared to modify its details in the light of later discoveries

363. But an important point finally emerges The cosmic dust-clouds appear in our treatment not as an accidental accompaniment of galaxies, but as an essential constituent of the scheme of the cosmos They are in process of building up, and have been in process of formation, by transit of particles across inter-galactic spaces, since the beginning of things They are still growing, and will continue to grow for ever For every particle we ourselves lose to other galaxies, we gain some other particle Indeed, as we have shown, if we lose (ultimately) a particle to galaxy A, we ultimately gain one from the same galaxy A Nebular evolution may thus be described, in relation to this phenomenon, as a continual interchange of partners On a gigantic scale, the dance of the universe is a chain-dance Dissolution is accompanied by re-formation, but re-formation of a new type of

system It is often asked, where could cast-off matter ultimately congregate? The answer is each galaxy is the rubbish heap of some portion of the universe Each galaxy, as it ages, loses distant members, but is recruited at the same time Any existing galactic neighbourhood is simultaneously the drill-ground of old soldiers and young recruits, but the young recruits are the emissaries of some other battalion

364. We ourselves lose members to a well-defined sub-set of galaxies defined by the sub-set of velocities $V'_0 = -ci \tanh 2\eta_l$, where η_l is a world-constant We gain our obscuring clouds from the same sub-set of galaxies These lie, at any instant in our experience, on a sphere centred at ourselves Each galaxy of the sub-set is likewise the centre of a new sphere of associated galaxies, with which it interchanges partners No two such spheres coincide Thus the cosmos possesses a sort of articulating organization, by which each nucleus of a sub-system is associated with a finite number of other sub-systems But the associated systems all overlap—each galaxy is a member of a number of associations The grand system of such intersecting and partially overlapping associations of sub-systems just occupies the expanding light-sphere centred at any arbitrary member of the system of nuclei, and this totality constitutes the universe Sub-systems—overlapping associations of sub-systems—the cosmos, such is our picture The interconnexion of galaxy with distant galaxy is a feature entirely unexpected when we embarked on our investigations, it gives a rational view of the course of nebular evolution

XIV

THE CAREER OF THE UNIVERSE

365. We have now completed an account of our main investigation The relation of the structure we have erected to the structures erected by the methods of general relativity will be considered in the next part We now propose to take a general view of the principles employed and results obtained

366. The systems of particles in motion which have been constructed must be carefully distinguished from their possible applications to the universe Whether they *represent* the universe is a matter partly for observation, partly for general considerations lying outside the scope of relativity pure and simple Thus whether the nebular nuclei are in uniform relative motion, whether their density-distribution is locally homogeneous but ultimately increasing outwards, whether there is an upper bound to observed distances, whether the average sub-system consists of outwardly accelerated particles, etc, are questions to be settled by observation Whether the universe contains an infinite number of systems can never be answered by observation in the affirmative, for we can at most count a finite number I have already given general considerations which lead us to conclude that the universe must include an infinite number of particles, but bearing in mind the historic vulnerability of general considerations I only point out here that whilst observation could conceivably verify the existence of a finite number of objects in the universe it could never conceivably verify the existence of an infinite number The philosopher may take comfort from the fact that, in spite of the much vaunted sway and dominance of pure observation and experiment in ordinary physics, world-physics propounds questions of an objective, non-metaphysical character which cannot be answered by observation but must be answered, if at all, by pure reason, natural philosophy is something bigger than the totality of conceivable observations But quite apart from questions of pure observation, and from the larger world-questions that cannot be settled by observation, our various systems of matter-in-motion possess what may be called an abstract reality, akin to any other abstract situation envisaged in natural philosophy

367. Theoretical physics consists essentially in the propounding of situations not existing in nature, and then in applying them to nature. Newton gave us superb examples. He introduced the concept of the particle acted on by no forces, which he stated would travel in a straight line with uniform velocity. When this statement has been dissected into definition and content, there still remains the circumstance that the situation it contemplates is never realized in nature. This does not in the slightest degree militate against the utility of the concept, or against its practical application to situations occurring in nature. A better example for our purpose is the two-body problem consisting of a massive gravitating body with a particle in motion in its presence. Newton from his dynamics and law of gravitation determined the nature of the possible motions of the particle and showed that it consisted of Keplerian orbits, or parabolas or hyperbolas with Keplerian properties. In nature we never have either two bodies existing 'alone', or, of the two bodies, one so small that it realizes the concept 'particle'. But this does not impair the 'validity' of the analysis of the abstract gravitational situation consisting of one heavy particle and one non-massive one.

368. The simple kinematic systems of equivalent particles, and the more complicated statistical systems, here developed, exhibit abstract gravitational situations of this kind. Whether they are realized in nature is an absorbing question, but one entirely irrelevant to their 'validity', just as Newton's solution of the Keplerian problem is independent of the degree to which it is represented in nature. The 'validity' or 'truth' of our kinematic systems is a question of the correctness of the arguments used and of the legitimacy of the ideal experiments on which they were based. I submit that they have an interest independent of the extent to which they reproduce observable characteristics of the universe.

I propose now first to review the abstract characteristics of the systems constructed, then to recount summarily their observable characteristics.

Abstract characteristics

369. The procedure has been to devise descriptions of events and phenomena capable of being made by a single observer with his own instruments, then to find the conditions to be satisfied by certain sets of phenomena when the descriptions by different observers are

compared. The instruments supposed to be in the possession of the observer are the clock, or time-measurer, and the theodolite, or direction-measurer. Every single formula we have stated has an immediate content in terms of clock-measures, theodolite-measures, and Doppler shifts. Clock-measures relate to events at the observer, theodolite-measures to directions recorded by the observer, Doppler effects to the ratios of time-differentials. An epoch t of an event *at the observer* is an immediate judgement. An 'epoch' t of an event *not at the observer* is the arithmetic mean of two epochs of events at the observer connected with the distant event by an ideal experiment consisting of the dispatch and return of a signal. A 'distance' r or $(x^2+y^2+z^2)^{\frac{1}{2}}$ of a distant event is the difference between the same two epochs of events at the observer, multiplied by an arbitrary constant $\frac{1}{2}c$. A coordinate x is the product of a 'distance' r from the observer and a direction-cosine, namely the cosine of the angle between some fixed direction and the direction of the observed signal received at the observer from the event, i.e. its apparent direction as seen in a theodolite. Actually we established the three-dimensional form of the Lorentz formulae by ideal experiments involving in effect parallax measures, for we used two distinct moving observers who had already determined their relative motion. But the diagrams we drew depicting the separate pictures made by the two observers could have been constructed equally from theodolite observations of directions by using the same Pythagorean theorem employed to combine the parallax observations. Thus a set of coordinates (x, y, z, t) is equivalent to a pair of time-measures at the observer and a pair of angles at the observer. When we say that each observer adopts Euclidian space for his own descriptions of events we merely mean that he has constructed coordinates out of observations made at himself by the above rules. In general relativity, on the contrary, events are sometimes described by means of coordinates constructed partly out of one observer's observations, partly out of another's. In our scheme the set of coordinates of an event is always definable in terms of observations made by the observer using these coordinates.

370. The observer then *defined* the velocity of a particle in his experience as $dx/dt, dy/dt, dz/dt$, where (x, y, z, t), $(x+dx, \ldots, t+dt)$ were neighbouring events at the particle. A velocity, like the position- and epoch-coordinates, is a construct by an observer out of his own

measures, and so relates to his own experience By an obvious extension of meaning of the word velocity, the signal velocity was found to be simply c, in the observer's experience A coordinate velocity can be immediately reduced to an observable Doppler effect.

371. Questions of relativity only arose when we attempted to compare the experiences of different particle-observers The general question was to find the relations between the coordinates (x, y, z, t) attached by A (from his measures) to an event P and those (x', y', z', t') attached by B (from B's measures) to the same event P when A and B could each (*a*) make the necessary ideal observations on P, (*b*) make observations on one another, it being understood between them that coordinates were to be constructed out of observations by the same formal rules, in particular by the selection (by agreement) of the same arbitrary $\frac{1}{2}c$ for assigning distances We tackled this problem in stages

372. We first considered the case of A and B making observations on one another We did not investigate the general problem of what restriction a knowledge of A's observations on B imposes on the possibilities of B's observations on A Instead, for historical and other reasons, we assumed the possibility of realizing in nature a pair of observers A and B such that A describes the totality of his observations on B in the same way as B decribes the totality of *his* observations on A By 'the possibility of realizing in nature' this situation I did not of course mean that such situations actually exist in nature any more than Newton assumed that there exist in nature 'bodies under no forces' I mean that it was legitimate to suppose that an abstract situation of this character was a possible situation in the sense of natural philosophy The *fact* of this being possible is our basis in experience of what follows Of course the situation has not been verified by experiments and observations actually carried out, but then no more has Einstein's situation that the velocity of light appears the same to two observers in uniform relative motion We are actually replacing the totality of experiences recorded (including such fundamental ones as the Michelson-Morley experiment) by a simpler experience which we believe to represent a possible situation, an experience which could be confirmed by actual measures capable in principle of being performed The situation we contemplated replaced the situation contemplated in

Einstein's axiom Einstein's axiom was wholly insufficient for us, because in mentioning velocity it already presupposed some unspecified scheme of assignment of epochs and distances to distant events, and its 'distances' involve the use of the indefinable concept of the rigid length-measure Our point of progress was that we provided actual tests, which A and B could carry out with their clocks, which would result in A and B being able to confirm or disprove whether they could describe their observations on one another in congruent terms They could thus ascertain whether or not they were 'equivalent'

373 We showed that when we added to the ideal experiments above mentioned the further ideal experiment of A and B reading *one another's* clocks, we were able to show that the conditions of equivalence implied the behaviours of their clocks in one another's experience This was the fundamental step The further steps of correlating the experiences of A and B in regard to any event whatever, namely the relating of (x, y, z, t) with (x', y', z', t'), were little more than corollaries from the clock-behaviour theorem We were led to transformation formulae of which particular cases, when the equivalent observers described one another as in *uniform* relative motion, reduced to the well-known Lorentz formulae The Lorentz formulae were thus derived from time measures only, without use of the indefinable concept of the rigid length-measure and without use of the axiom of the constancy of light-velocity to observers in uniform relative motion, simply from the axiom of the possibility of the existence of equivalent observers, which is much wider than any axiom relating to uniformly moving observers

374. The upshot of this situation is that each of a set of equivalent observers in uniform relative motion can employ, for reducing his own measures, Euclidian space and Newtonian time, and make diagrams and calculations accordingly The Lorentz formulae effect a correspondence between the various Newtonian times and Euclidian spaces adopted by the various equivalent observers The resulting physics, in the experience of each separate observer of the set of equivalent ones, is simply ordinary classical physics

375. The contrast between our procedure and that of 'general' relativity lies partly in the circumstance that we have begun with actual

observations and constructed coordinates and relations between coordinates out of them 'general' relativity begins with a conceptual geometry in which events are described by conceptual coordinates, and later (by means of the definition of a light ray as $ds = 0$, $\delta \int ds = 0$) ascertains the observations implied by this conceptual scheme 'General' relativity is concerned with the class of relationships, and the class of systems of particles in motion, describable by means of a Riemannian metric, it never claims to show that all possible systems of particles in motion can be so described Consequently it is not remarkable that at least some systems are incapable of description in terms of its conceptual scheme On the other hand, our method ensures the possibility of descriptions of any system containing equivalent observers by means of observations made by those observers, these systems, as will be shown, include systems not described in 'general' relativity

376. The next difference from 'general' relativity arose in the use of the application of the principle of relativity itself 'General' relativity assumes the proposition that a relation describing a physical law must be expressible in the form of the vanishing of some quantity in all systems of coordinates But it makes no distinction between transformations of coordinates which are mere transformations of observations made by a single observer *inter se*, and transformations of coordinates which involve the mixing of observations made by one observer with observations made by one or more other observers Transformations of the former type are of a merely mathematical content and add nothing from the point of view of natural philosophy, they are tautomerisms, alternative ways of expressing one observer's observations in symbols That the expression of a law of nature in the language of 'general' relativity includes the totality of such special transformations is the real reason for the cumbrousness of the tensor calculus, it is as though a language were being used which contained all possible languages, all possible syntaxes, as particular cases But more important still is the consideration that 'general' relativity does not clearly distinguish between transformations which are tautomerisms and transformations which contain a factual content, namely transformations which relate the experiences of one observer to the experiences of a second observer Such transformations add to the content of an enunciation, from the point of view of

natural philosophy Stripped to their essentials, the transformations which alone have a natural-philosophical content connect observations made by one observer at himself (for observations are only possible by an *ego*) with observations that would be made by another observer at *himself*, when the motion of the second observer in the experience of the first observer is specified by relations defined by observations made (or capable of being made) by the first observer on the second observer—each observer being idealized as a particle Thus the transformation must connect straightforward clock-and-theodolite observations, a clock-plus-theodolite being carried by each separate observer

377. Our method on the contrary was to transform not from any set of coordinates to any other but to transform from any one observer to any other observer For this purpose it is sufficient to use only one set of coordinates, and that the simplest, for each observer, we chose for him Cartesian coordinates and Newtonian time We actually found it necessary, in our application of the method, only to transform from one observer to another *equivalent* observer But this must not be allowed to obscure the essential nature of the method, which is presumably capable of much wider application

378. As a consequence of the method we were able to apply the principle of relativity in a new way, to determine the form of what may be called a 'law of nature' on *a priori* grounds This became possible when we constructed a system of moving particles such that they not only were equivalent in pairs, but possessed equivalent descriptions as a whole by any two (equivalent) particle-members We described such a system as satisfying the cosmological principle, and defined it by the condition if two particle members A and B are such that $A \equiv B$, then also $A \equiv B$ We then inquired as to the description of the acceleration of an additional freely moving test-particle added to the system Since all the given particles are equivalent, and describe the whole system equivalently, the description of the acceleration by means of a function of arguments defining the freely moving particle must be *of the same functional form* in the experience of each of the given particle-members of the system, using coordinates constructed out of observations by the same rules of combination There are two essential conditions satisfied which ensure the legitimacy of this application of the principle of relativity,

namely that the observers enumerating the acceleration function are equivalent and that they adopt what may be called *congruent* co-ordinate systems It would be no use for one observer to use Cartesian coordinates and another spherical polar coordinates It was for this reason that we had first to begin with observations, then to prescribe identical rules for the observers to use in compounding their observations to yield coordinates To begin with a conceptual scheme of unspecified coordinates, as in general relativity, would make the whole procedure inapplicable Thus in our line of thought a 'law of nature' satisfied a condition of conservation of *form* (of a function), not as in general relativity a condition of conservation of values (of a tensor). General relativity is like a garden where flowers and weeds grow together The useless weeds are cut with the desired flowers, and separated later In our garden we try to cultivate only flowers.

379. We were then able to trace the general properties of trajectories of freely moving particles, and to show that the originally given particles, of prescribed motions, could also be considered as freely moving, free from constraints (A similar procedure occurs in general relativity, where, when a metric is chosen, the trajectories are defined as the geodesics, and it is later shown that the originally prescribed particles-in-motion whose constrained motion gives rise to the 'field' described by the metric will in fact follow free paths) We thus obtained a great deal of information about certain systems of given particles in motion and the possible motion of an additional free particle at large in the system We can call such a system a 'gravitating' system if we wish, though we gain nothing by using the word 'gravitating' The important point is that we thus construct, in the flat space and Newtonian time of each observer, a 'natural' system of non-zero particle-density moving without constraints, without any recourse to the 'curvature of space' on the one hand, or to any specific 'theory of gravitation' or 'field equations' on the other hand No physical constants were necessary to its description, the one arbitrary constant B, a constant of proportionality in the density-distribution, simply defined a continuum of possible distributions

380 The system was completely 'relativistic' in the only significant sense of the word, namely that its members, describing their experiences according to the same rules for combining observations, had identical experiences—or rather superposable experiences.

381. The construction, in even one particular class of instances, of a gravitating system without recourse to the curvature of space or any specific theory of gravitation constitutes an advance on the so-called 'general' theory of gravitation or relativity The latter theory adopts, quasi-empirically, a 'law of gravitation' or 'field-equation' which embodies the invariant form of the Newtonian scheme of gravitation and dynamics, but it gives no reason why this particular form should be chosen, and indeed explicitly recognizes that other field-equations, containing an arbitrary constant λ, would be equally legitimate, nor does the introduction of a 'curvature of space' go any distance towards 'explaining' gravitation, for it leaves unanswered the questions 'what is the thing that is curved?' and 'why is it curved?' The general theory of gravitation piously hopes that nature favours some particular value of its arbitrary constant λ, but it is bound to confess that as far as its own principles go all values of λ are equally acceptable Eddington has indeed attempted to face this question, but his answer has not won general acceptance Our system, on the other hand, contains no physical constants whatever, either arbitrary to the theory or special to nature The only constant occurring is c, which was an arbitrary original choice on the part of one observer, communicated to other observers †

382. This suggests the inspiring possibility that all 'gravitational' situations ought to be capable of similar treatment Only future research can show this, but it is part of my faith that all gravitational motion, just as in the world-systems we have constructed, is but the only possible totality of motions appertaining to a given prescribed abstract situation which are capable of consistent description by all the particle-observers present, and that such description will contain no 'natural constant' whatever It is quite clear that, once we have constructed a single gravitating system free from introduced natural constants, to take any other view of gravitation would be a retrograde step The general progress of physical science lies in the way of reducing more and more experiences to fewer and fewer principles, and of reducing more and more constants to combinations of a small number of constants In the world-systems we have constructed we have carried this aim to its logical conclusion—we have expressed

† I refer here to the simple kinematic system The more general system contains a constant of integration C, but arguments already given are in fact equivalent to the fixing of C in terms of the arbitrary function $\psi(\xi)$ defining the statistical system

the gravitational system in question in terms of no irreducible constants whatever The 'cosmical constant' λ is thus most probably only an accident due to the particular mode of expression adopted by general relativity Instead of introducing this constant by an arbitrary assumption and then trying to get rid of it by reducing it to a combination of other constants, as Eddington does, it is surely better to aim at an enunciation of gravitation which itself contains no constants and which therefore assumes neither the existence nor the non-existence of the 'cosmical constant'

It is scarcely profitable to attempt to forecast the future of research on gravitational problems on these lines, but some indications may be given Suppose the problem is the Keplerian problem of the test-particle P in motion round a massive one O Then the totality of the observations made by P on O must be describable by P in the same functional form as the totality of observations from any other test-particle P' on O, made by P' Transformation from P to P' will provide a condition to be satisfied Since the 'gravitational field' surrounding O is simply the summary of the motions of all possible test-particles, if we can describe the totality of observations from O of test-particles P, or of observations from all P's of O, we shall have found the utmost of which analysis of the system is capable We may expect in this way to obtain simultaneously the 'laws of motion' and the 'field' A similar method could in principle be applied to any local gravitational situation The technical difficulty is to find the transformation from P to P' But in all cases the significant step would be the transformation, not of coordinates, but from observer to observer when each observer uses congruent coordinates, i e coordinates constructed out of observations by the same rules It is the transfer of observer's head-quarters which matters

383. When, in our programme, we added not only one free particle to our simple kinematic system, but a multitude, and arranged the new system to satisfy the cosmological principle, we were able to determine fully the 'gravitational field'—we were able to extract the utmost possible from the principle of relativity in our form This 'utmost possible' proved to be precisely the complete description of all the motions, given the particles present given ψ, which defined the statistical description of the population of particles present, we obtained a description of the trajectories followed, in the minutest

detail, and that for any ψ Thus we accomplished for these systems a complete gravitational programme without recourse to 'field equations' or to any specific formulation of a 'theory' of gravitation The results are thus of an absolute character They should be compatible with any *specific* theory of gravitation which is strictly relativistic But they avoid the ambiguity present in Einstein's law of gravitation—the ambiguity in density, given a metric, due to ignorance of what value to attach to 'λ' It is irrelevant to plead that λ may be small, the ambiguity is definitely present

384 We could, if we wished, analyse our changes of motion in terms of action at a distance We should, again, be at liberty to consider such actions at a distance as propagated instantaneously, or as propagated with the speed of light, the results of the analysis would be different in the two cases For example, we might find that on one mode of analysis we could introduce a constant parameter γ which we could call a constant of gravitation, on other modes of analysis we might find the same parameter to be effectively dependent on the time, an effect which might be described as arising from the 'influences' of distant particles retarded by a finite time of propagation Actually the analysis of our statistical system contains only the constants C and η_l, both connected with the function ψ defining the distribution considered

385 With regard to our avoidance of the introduction of the 'curvature of space' or of 'space-time', it is, I think, becoming gradually better recognized that the phrase 'the curvature of space' denotes nothing physical in nature, that the phrase 'physical space' is meaningless, and that to use a metric defining a Riemannian continuum is but one amongst many possible choices of mathematical procedure useful for describing certain classes of phenomena Any notion that bodies are accelerated 'because space is curved' or that nebulae separate 'because space is expanding' is of course meaningless Space and space-time are constructs, manufactured for convenience, with no more physical content than the phrase 'the ether' We have made our work intelligible by employing, of all the spaces at our legitimate disposal, the space commonly used in physics This differentiates our treatment markedly from the modes of expression used in current gravitational relativity We have employed for the time of each observer the time commonly used in physics, the time by which

velocities are measured, races timed, and railway time-tables constructed This has for a consequence that all false paradoxes and contradictions with experience are avoided In our work it would be impossible for two particle-observers to meet with synchronized clocks and meet later to find the one younger than the other No meaning can be attached to an observer's 'waving his now' about in a continuum

Observable characteristics

386 The observable characteristics of the simple kinematic system of equivalent particles with a hydrodynamic type of flow have been summarized in Chapter VI I propose here to summarize the characteristics of the statistical systems which have been constructed by adding other free particles to the simple system until the cosmological principle is again satisfied These abstract gravitational situations have the following properties

(1) the trajectories divide themselves into sub-systems,

(2) the particles, members of any one sub-system, are concentrated towards a nucleus,

(3) the nuclei follow the motions of the 'simple' kinematic system

  (*a*) they recede from one another,

  (*b*) their motions satisfy a velocity-distance proportionality at any one epoch in the experience of any one of them,

  (*c*) their velocities are constant in the experience of any one of them,

  (*d*) they possess no preferential velocity-frame,

  (*e*) each nucleus is central with regard to all the others,

  (*f*) the nuclei are distributed approximately homogeneously in the immediate neighbourhood of any one of them,

  (*g*) the density of nuclei increases at great distances from any one of them, tending to infinity at distance ct, where t is the present epoch,

(4) the particles, members of a sub-system, are distributed in density round their nucleus according to a law of the inverse cube, or a law somewhat more severe,

(5) the totality of sub-systems is contained within an expanding sphere of radius ct centred at any nucleus, where t is the present epoch reckoned by the observer on the nucleus, the boundary of this sphere is inaccessible,

(6) the members of any sub-system are in outward accelerated motion from their nucleus,

(7) the members of the different sub-systems are in transit across inter-nuclear space and intermingle,

(8) the nuclei determine a background of finite luminosity which is never ultimately resolvable,

(9) the total number of sub-systems is infinite,

(10) every particle attains the speed of light at a finite epoch and at a finite distance in the experience of any assigned nuclear observer,

(11) in any arbitrary volume of space there are present at any epoch particles moving with speeds in the vicinity of that of light,

(12) after attaining the speed of light each particle decelerates and attains a constant limiting velocity as $t \to \infty$, in the experience of any nuclear observer,

(13) in the vicinity of each nucleus there ultimately arrive particles at rest relative to this nucleus, from other sub-systems,

(14) for each particle which the sub-system belonging to a given nucleus ultimately yields up to another nucleus, the given nucleus ultimately receives a particle from the sub-system belonging to the second nucleus,

(15) each nucleus is associated with a definite set of other nuclei, to which it sends, and from which it receives, ultimately uniformly moving particles

387 Overlooking local deviations from pattern, we have identified the system so described with the universe described in astronomy We begin by identifying the sub-systems with the extra-galactic nebulae and the sub-system nuclei with the nebular nuclei This may be called our principal identification hypothesis, and covers properties (1) and (2) Any further reproduction of the properties of the abstract system by the observed universe may be regarded as confirmation of this hypothesis

388. The possession by the nebular nuclei of the properties enumerated under (3) has been the subject of Chapter VI The recession and recession law confirms (3) (*a*) and (3) (*b*) Characteristic (3) (*c*) is not yet verifiable within the present range of observation, but may soon be tested when regard is paid to the fact that in a photograph or

spectrogram the positions and velocities recorded correspond to different epochs for objects at different distances Characteristic (3) (*d*) is demanded by our belief that the universe will possess no preferential frame of reference Characteristic (3) (*e*) is assigned to the universe on almost all cosmologies Characteristic (3) (*f*) is confirmed by Mount Wilson and Harvard surveys, though some recent work by Shapley calls it in question Characteristic (3) (*g*) is scarcely within present observable range

389 The law (4) of concentration of matter towards a nucleus, together with the spherical symmetry of the sub-systems when undisturbed by tidal effects or other local irregularities, is compatible with the observed luminosity-distribution in nebulae, its rapid increase towards an almost-point-nucleus, and the circular shape of the observed isophotic contours near a nebular nucleus

390 Characteristic (5) is not within the range of present observational verification But if Hubble's law continues to hold—and its linear form is an essential consequence of hydrodynamical continuity—some upper limit to the distances must exist, as otherwise velocities exceeding the light-velocity would be predicted

391 Characteristic (6) is confirmed by the observed K-effect inside our own galaxy, and by the outward motions suggested by the observed forms of many nebulae It is true that the existence of the K-effect in its usual form, in our galaxy, has been called in question by recent work by J S Plaskett, but it must be remembered that an outward motion from the galactic centre will only reveal itself locally, at an excentric point like the sun, by a differential K-effect increasing with distance and different in different directions, and observed radial velocities have not yet been analysed for a differential radial component of this kind Such an effect may be present, but not yet disentangled from the differential galactic rotation effect of Oort, which would partly mask it It is to be hoped that analysis of stellar radial velocities for a differential K-effect varying with galactic longitude and latitude but symmetrical about the line joining the sun to the galactic centre will soon be undertaken

392. Characteristic (7) is in accordance with evidence recently adduced by Larmor as to the motion of stars and star-clusters into our own system from outside

393. Characteristic (8) is in accordance with present observation An ultimately resolvable background would create tremendous philosophical difficulties, as we shall see in Part IV

394. Characteristic (9) can never be confirmed by observation But no limit to the number of observable nebulae has yet come into observational recognition, and we saw in Chapter VI that a universe of a finite number of objects would again create profound difficulties

395. Characteristics (10) and (11) are in accordance with the phenomena of cosmic rays, more especially their absence of association with particular directions, the corpuscular character of the primary agent, and possession by such primary corpuscles of speeds indefinitely close to that of light

396. Characteristics (12) and (13) are in accordance with the observed existence of cosmic dust-clouds and other obscuring clouds, travelling with their respective galaxies Characteristics (14) and (15) give a rational picture of the development and history of these clouds

397 The observed characteristics of nebular proper motions, of nebular flattenings and rotations, and of the spiral forms of nebular arms are not formally accounted for in the ideal scheme we have described But place for them exists in the scheme For the nebulae are discrete, finitely separated, systems, whilst our analysis has throughout been statistical in character, treating of elements of volume large enough to contain a number of units and so covering only *average* properties of sub-systems The random orientations of the observed planes of spiral nebulae is in accordance with our analysis But a more finely detailed analysis would endeavour to replace the continuous statistical calculus by a calculus of finite differences If such a calculus is attempted, it appears impossible to construct an abstract system of discretely separated particles satisfying the cosmological principle in three dimensions † This is an important negative result It is therefore impossible to approximate fully to the statistical system by a system of discretely separated objects, though such a system can be constructed easily in one dimension A discretely separated set of particles may be arranged so as to satisfy statistically (i e on the average) the cosmological principle, but the detailed aspects of the system as viewed by each member must differ

† See Appendix, Note 8

slightly from member to member A most fascinating domain of mathematics opens up at this point, for we could investigate the condition of minimum deviation from satisfaction of the cosmological principle for a set of finitely separated particles I do not propose to examine this problem here, but it is clear that differences of aspect will be associated with residual accelerations, fluctuations, or deviations from the smoothed-out values, which in turn may set up local proper motions, rotations, tidal flattenings, and other features of form and motion It is not improbable that such dynamical consequences can be inferred by purely kinematic methods, for example, the departure of the grand system from perfect spherical symmetry round each individual member will be expressible as the coefficients of a momental quadric, whose axes will determine components of acceleration and associated couples Thus whilst statistical analysis gives the broad features of the world-system, individual peculiarities of form and motion must await a more powerful and more detailed technique capable of dealing with the 'fine structure' of the units concerned It would, indeed, have been wholly unsatisfactory if our crude statistical analysis had predicted the forms of individual nebulae, we may regard it as paving the way for a calculus which, taking due account of finiteness of separation of nuclei and the three-dimensionality of space, will examine subordinate details Our present calculus has played the important part of indicating which are merely fine-structure details and which are large-scale features

398 One point on which criticism may be anticipated may be mentioned Each sub-system in our scheme contains an infinite number of particles, the infinity occurring at the nucleus of the sub-system In one way this infinity in the ideal scheme is satisfactory it is the way chosen by the mathematics to define a strong central nuclear condensation, and to ensure the existence of this condensation for all time, for a finite population for each sub-system, owing to its centrifugal expansion, would eventually be impoverished to zero But it may be asked what 'influence' the infinity in density at each nucleus has on the forms of trajectories This question can be answered in general terms, and the singularity in density removed, by subtracting from the statistical system a simple kinematic system corresponding to the nuclear singularities This will leave the motion of each nucleus unaltered—its velocity will continue constant—since

the 'field' of the simple kinematic system reduces to zero at each fundamental particle. The details of the forms of individual trajectories of other free particles will be affected, but not their general features, for the acceleration-formula will remain of the same functional form We cannot simply superpose accelerations, of course A new analysis is required, combining the hydrodynamical and statistical treatments, much in the same way as mixed radiation fields, consisting of continuous pencils of radiation and parallel beams of radiation, can be treated in theoretical optics I do not here attempt this analysis, but the idea of superposing a hydrodynamical system (taken negatively) on a statistical system may prove a fruitful one *A priori* it offers the possibility of reducing the population of a sub-system to a finite number

399. Our treatment has throughout been guided by the desire not merely to express the universe in symbols but to go behind the symbols to the phenomena It is thus in marked contrast not only with current relativistic cosmology but also with much of modern theoretical physics, which denies that symbols can be used to give insight into the relations between phenomena

The career of the universe

400. We have described a possible career for the universe from the dawn of time to the everlasting future, from 'creation' to 'world without end' In our abstract idealization we trace the history of each individual constituent from $t = 0$ to $t \to \infty$ in the experience of any of the fundamental equivalent observers, and we describe the relationship of each particle to other particles, their memberships of sub-systems, the events along their trajectories and their ultimate fate No such comprehensive scheme of motions, no detailed treatment of the infinitely-many-body problem, no complete world-history, has been previously sketched

401. We briefly restate this world-history The epoch $t = 0$ defines an initial singularity It is a singularity in total density, and a singular point on each trajectory From this event each particle starts with some definite initial velocity $\mathbf{V}_0$, where $0 < |\mathbf{V}_0| < c$ and $\mathbf{V}_0 = 0$ specifies the arbitrary particle-observer O in whose experience the initial velocity was $\mathbf{V}_0$. But since the point $t = 0$ is a singular point on the trajectory, the value of $\mathbf{V}_0$ does not fully specify the trajectory The trajectory is defined by two other constants, a vector $\mathbf{i}$ and

a scalar X_l. The particles with a common initial velocity $\mathbf{V}_0$ form a sub-system. The 'centre' of the sub-system moves with the constant velocity $\mathbf{V}_0$, and follows a singular trajectory, namely a straight line. The members of the sub-system separate from one another according to the values of $\mathbf{i}$ and X_l, but move with the average velocity $\mathbf{V}_0$. Seen from its centre, each sub-system possesses spherical symmetry, each member moving radially outwards from the centre in a way which depends on the whole statistical distribution of particles present. Each member has for its hodograph (in O's experience) a straight line defined by $\mathbf{V}_0$ and $\mathbf{i}$, the different members with the same hodograph differing in their X_l's. In the experience of O, who is the centre of the sub-system $\mathbf{V}_0 = 0$, a member of the $\mathbf{V}_0$-sub-system ($\mathbf{V}_0 \neq 0$) may be at first accelerated or decelerated, but it is always eventually accelerated. This acceleration (or deceleration) is always directed towards the apparent centre of the whole system to the particle concerned, namely the position of that 'fundamental' particle $\mathbf{V}_0 =$ constant for whom the constant is equal to the instantaneous velocity $\mathbf{V}$ of the particle concerned, all reckoned in O's experience, in the frame of this fundamental observer $\mathbf{V}_0 = \mathbf{V}$, the particle concerned is momentarily at rest. (Each fundamental observer is not only the centre of his associated sub-system, but also central, in his own experience, with respect to the totality of particles present.) As $\mathbf{V}$, the velocity of the moving particle concerned, changes, the fundamental particle relative to whom the moving particle is instantaneously at rest alters, and the apparent centre of the universe (to the moving particle) migrates from fundamental particle to fundamental particle. It has been shown that though always 'falling freely' towards the apparent centre (to it) of the grand system, the moving particle never reaches the migrating apparent centre, save possibly as $t \to \infty$. Instead, in its 'endeavour' to reach this apparent centre, it accelerates until it attains the velocity of light, which it does at a finite epoch and at a finite position interior to the whole system. Whilst swiftly moving, with a velocity in the neighbourhood of that of light, it may take part in a collision and give rise to a cosmic-ray phenomenon. (It may, indeed, undergo collision before or after this juncture, in which case it may take part in a process of dust-particle-building or star-building.) After attaining the velocity of light, if it continues to avoid any collision, it decelerates and, as $t \to \infty$, tends to assume a constant limiting velocity.

(The acquisition of the speed of light and the subsequent deceleration may be described as due to the 'gravitational pull' of the whole universe of matter in motion But it will be noticed that the notion of a gravitational potential is completely inapplicable to the system, partly because it is impossible to get to 'great distances from all attracting matter' (the usual dynamical phrase), partly because the system possesses no 'field of force' in the sense in which the phrase is used in discussing conservative systems in dynamics Though the acceleration formula explicitly involves the velocity it is not to be imagined that the acceleration 'depends on' the velocity, the velocity has to be mentioned simply because an observer has to be mentioned in whose experience the motion is being described, but the acceleration 'depends' only on the distance of the particle concerned from the apparent centre of the system in the frame in which it is at rest)

The members of a sub-system are strongly concentrated towards the centre of the sub-system, which thus forms a nucleus These nuclei we identify with the nuclei of the extra-galactic nebulae, and we accordingly identify the sub-systems with the nebulae themselves The members of the different sub-systems intermingle, any assigned volume of space containing a single nucleus will contain, in addition to the particles belonging to this nucleus (possessing the same value of V_0), particles belonging to other, distant or neighbouring, nuclei Internuclear space will accordingly be peopled by 'field-particles', derived from a variety of nuclei and belonging to a variety of sub-systems Near an assigned nucleus, the field-particles or intermingling members from other systems will be in a negligible minority compared with particles 'belonging to' the nucleus concerned These latter particles stream outwards from the nucleus, their outward motions constituting what is usually described as a K-effect, and helping to give rise to the observed forms of the arms of spiral nebulae.

The members of each sub-system appear to have originated from the associated nucleus But they may be considered as having equally originated from any other nucleus, for all issued from the primeval singularity at $t = 0$ They 'belong to' a nucleus simply in the technical sense of possessing the same integration constant $\mathbf{V}_0$ But their density-concentration allows immediate recognition of the value of V_0 characterizing them, namely the V_0 or constant velocity of the nucleus The separation of the nuclei carries with it the partial separation of the sub-systems, in the sense of the separation of the

majority of the members of a sub-system from the majority of the members of any other sub-system.

Each nuclear region steadily loses members. Theoretically, in the abstract scheme, the process never terminates, but goes on for ever without the density-concentration ever disappearing. For the density-concentration is represented in the ideal scheme by an infinity in particle number. Actually, for the finite-density concentrations of nature, the process must result in the complete dissolution of the sub-system. But the nuclear vicinity does not go uncompensated. Each nucleus gathers in its vicinity particles arriving at relative rest from other sub-systems, shed off as it were from other nuclei. These arriving particles ultimately move along with the nucleus which is their destination. They may be identified primarily with the clouds of obscuring matter found in association with galaxies. Such dust-particles may be formed from the coalescing of atoms, electrons, positrons, or other primary corpuscles represented by our ideal particles. But larger formations, such as stars, multiple stars, planetary systems, or star-clusters, may also ultimately arrive and build up a new galactic system. Thus every sub-system is in process partly of continuous dissolution, partly of continuous re-formation, the new arrivals may ultimately build up a complicated, uniformly moving cosmic structure.

402. It might be thought from this, at first sight, that the *ultimate* state of the universe is a collection of moving cosmic clouds and other formations replacing the original galaxies. This is not so. The universe possesses no ultimate state. Each individual galaxy disperses, and is replaced by an obscuring cloud or other aggregation, as $t \to \infty$, as reckoned in its own local time-history, $0 < t < \infty$. But at great distances from it, in its own time-experience, the galaxies are at a much earlier stage of evolution. That is to say, at events on the V_0-galaxy at the world-wide instant t in the experience of an observer at the nucleus of a given galaxy $V_0 = 0$, the observer on the V_0-galaxy will assign to these events the earlier epoch $t(1-V_0^2/c^2)^{\frac{1}{2}}$, and therefore his stage of dissolution and re-formation will be less advanced. Ever, on the confines of the observable universe, galaxies could in principle be observed in arbitrarily early stages of evolution, in a stage arbitrarily near $t = 0$, i.e. arbitrarily near 'creation'. Thus though the process of dissolution and re-formation is ever in being for each

individual sub-system, we can never say that it is nearing completion for the totality of galaxies The evolutionary process, however far advanced at any individual nebula, has proceeded along only an arbitrarily short course for other sufficiently distant, sufficiently swiftly moving galaxies in the world-wide 'present' of the given individual nebula The universe goes on for ever The evolutionary process never terminates, it never even approaches termination We can no more speak of the 'age' of the universe, objectively, than of the 'radius' of the universe, each refers to a particular observer at a particular stage of his experiences To us, now, at epoch t measured from the natural time-zero, the 'age' is t and the 'radius' ct, but at an event in our world-wide 'present', t, for the observer experiencing this event on the nucleus moving with velocity $\mathbf{V}_0$, the age is $t(1-V_0^2/c^2)^{\frac{1}{2}}$ and the radius $ct(1-V_0^2/c^2)^{\frac{1}{2}}$, and thus both are smaller

Thus whilst each individual sub-system has an ultimate fate, the universe as a whole has no ultimate fate Referring to the whole universe, the word 'ultimate' has no meaning All statements to the effect that the universe ultimately stagnates, or passes to a state of 'heat-death', all attempts to invoke the second law of thermodynamics on the cosmic scale, are devoid of meaning Each separate, limited aggregate such as a star or galaxy may cool or decay, the second law may be validly applied to any finite system not comprising the totality of things, but to apply it to a universe containing an infinite number of entities has no meaning Entropy, as Eddington has remarked, points time's arrow, but time is rather to be regarded as a flight of arrows, each with its distinct course, and no entropy-maximum is ever attained by the universe, even as $t \to \infty$, for every experience, however large t may be, includes events for which the local time and so the local measure of the local entropy-increase are arbitrarily small The universe, if capable of representation by our ideal scheme, as it seems to be, is an ever-continuing system, knowing birth but not death Each limited portion, each nebular system, decays, perhaps dies, but the race of nebulae survives for ever

Into the further philosophical implications of these conclusions, with their frank contrast with the expressed views of writers like Eddington and Jeans, which have been taken so seriously by thinkers such as Inge, I shall not here enter I would only remark that previous discussions of the cosmological problem have scarcely passed beyond the rudimentary 'hydrodynamical' stage comparable

with our discussion in Part II, and that no complete statistical analysis of the universe as a collection of separate moving particles has been previously given. By re-analysing the foundations of relativity, without making use of any so-called 'laws of nature' or 'natural constants', we have given answers to questions which have engaged the attention of philosophers since fundamental inquiry began. I am deeply conscious of the many imperfections in the answers, and our attempt at a complete cosmology must necessarily be regarded as a provisional solution, liable to drastic revision or even complete rejection as investigation and observation advance. I merely submit that our answers to the questions of time and space, of past and future, of geometry and gravitation form a self-consistent whole, free from internal contradiction and resting on an extremely small substratum of empirically ascertained fact, this substratum comprising little more than the existence of a temporal experience for each individual.

403. Minkowski once said: 'The whole universe is seen to resolve itself into world lines, and I would fain anticipate myself by saying that in my opinion physical laws might find their most perfect expression as reciprocal relations between these world lines.' In a very imperfect way we have carried out Minkowski's programme; we have avoided making any hypotheses of a physical character about laws of nature, yet we have arrived at a complete description of an infinity of possible motions of systems which may be described as gravitating systems. Whether the distribution of matter-in-motion 'causes' the accelerations undergone by the particles we have never had to inquire; we have found it sufficient to ascertain what accelerations coexist with specified distributions of matter-in-motion. Just as an absolute time disappeared in Einstein's analysis, so an absolute law of gravitation has disappeared in ours. Minkowski's programme has proved completely practicable.

404. Kirchhoff once referred to 'the formulation of the natural sequence of the motions of inanimate material systems'. This is Minkowski's programme expressed in other words.

405. Sir Joseph Larmor has written: 'To say, as is sometimes done, that force is a mere figment of the imagination which is useful to describe the material changes that are going on around us in nature is to assume a scientific attitude that is appropriate for an intelligence

that surveys the totality of things ' Without claiming the possession of such an intelligence, we may assert that this is precisely the attitude we are compelled to adopt in attempting a solution of the cosmological problem 'Force' we have in fact regarded as a figment of the imagination 'Matter and motion', in Clerk Maxwell's phrase, are alone observable, counts of particles, clock-measures, Doppler effects, form the world-stuff we have discussed It is not surprising therefore that a purely kinematic treatment, avoiding dynamics, has proved adequate for our purpose, for our material is of kinematic character Dynamics is an intermediary between man and motion

406. Lord Kelvin once asked 'Were the primordial atoms relatively at rest in the most ancient time?' 'Or were they moving with velocities relative to fixed axes through the centre of inertia of the whole, sufficiently great to give any contribution to the present kinetic energy of the universe?' 'It is conceivable that all the atoms were relatively at rest in the most ancient time '

We have obtained an answer to Lord Kelvin's first question, and answered his conjecture in the negative The atoms were not at rest, for that would have been to give a meaning to 'rest' in contradiction with the impossibility of describing an absolute frame of reference Our proposed representation of the universe contains no frame of 'absolute rest' The system has no centre of inertia, only an apparent centre in the experience of any observer, and different centres for different observers

407. The writer of the article 'Herbert Spencer' in the 1911 edition of the *Encyclopaedia Britannica* remarks 'Spencer appeals alternately to the "instability of the homogeneous" and the impossibility of complete equilibration to keep up the cosmic see-saw, but he can only do so by confining himself to a *part* of the universe A world wholly homogeneous or equivalent could no longer change, whilst so long as a part only is in process of change, the process cannot be represented as universal Again an infinite world cannot be wholly engaged in evolution or in dissolution, so that it is really unmeaning to discuss the universality of the cosmic process until it is settled that we have a universe at all capable of being considered as a whole '

I submit that our provisional solution of the cosmological problem answers Spencer's arguments Though Spencer saw through a glass darkly, his arguments have their full force But since our inferred

world-structure has a finite volume at any one time (in the experience of any one observer) it *is* capable of being considered as a whole, though from the inside only. Change is provided for by the all-pervading evolutionary process, advanced near ourselves and only just begun near the confines. We meet Spencer's difficulties by making use of the mathematical concept of an *open* set of points, an infinity of points in a finite domain deprived of its limiting points so that every point is surrounded by others. There is no cosmic see-saw. Experience, the totality of all experiences, is a succession of overlapping and related life-times. Experience of the past is communicable, experience of the future unattainable. The inevitability of the local evolutionary decay-process is accompanied by a perpetual all-but-re-creation near the boundary, elusive, rainbow-like, incapable of capture. Spencer rightly urged the impossibility of the homogeneous and the static. Motion is nature's answer to problems of time, nature's device for constructing an infinitely populated universe in a finite space possessing, owing to the inaccessibility of its boundary, all the properties of infinite space but avoiding the difficulties of *an* infinite space.

408. To discuss time is to discuss the core of experience. To discuss space is to discuss the scaffolding of a building. The scaffolding is put in by the observer, is necessary for the examination of structural details. But the scaffolding is arbitrary. The resulting mode of description will depend on the scaffolding, will be a function of points of view. But what is important is what is inside the scaffolding, the structure itself.

PART IV

WORLD-PICTURES

XV

THE WORLD-PICTURE ON THE SIMPLE KINEMATIC MODEL

409. The object of Part IV is to compare the kinematic structure we have erected with Newtonian and with relativistic dynamics and their associated cosmologies

410. In the scheme of 'general' relativity any proposition represented by symbols is capable of immediate translation into observations Similarly in our kinematic scheme any proposition is immediately translatable into observations Therefore a complete comparison can be made between the relativistic cosmologies and our own by turning both into the implied observations and then comparing the observations As our view is that the geometry adopted by an observer is arbitrary, it would be idle to compare geometries, we should get no farther Likewise it would be idle to compare any propositions stated in terms of coordinates We must go back to the actual observations out of which the coordinates have in our systems been constructed, and the actual observations implied by the conceptual constructs of 'general' relativity

411 The plan of Part IV is that we first briefly collect together the formulae giving the observable properties of the kinematic structures We then construct a complete cosmology based on Newtonian dynamics and gravitation, within the domain of validity of the Newtonian system We next observe that the symbolic representation of the Newtonian cosmology is formally identical with the symbolic representation of the cosmologies of 'general' relativity, but that the latter representation requires a different interpretation in terms of observations This interpretation we proceed to obtain in a form strictly comparable with the interpretation of our simple kinematic models in terms of observations Comparison then shows that certain relativistic cosmologies under various circumstances approximate to our systems, and even tend to them in certain limiting cases, but differ

from them essentially in definite and important respects We come to the conclusion that the relativistic cosmologies must be rejected as contradicting experience, whilst the kinematic models are compatible with experience

Comparison with local Newtonian gravitation and dynamics

412 The simple kinematic model is described by the observer at the origin as possessing the particle-density distribution

$$n\,dxdydz = \frac{Bt\,dxdydz}{c^3\{t^2-(x^2+y^2+z^2)/c^2\}^2}, \tag{1}$$

where B is arbitrary, as satisfying the velocity law

$$u = x/t, \qquad v = y/t, \qquad w = z/t \tag{2}$$

and as possessing the velocity distribution law

$$n\,dxdydz = n'\,dudvdw = \frac{B\,dudvdw}{c^3\{1-(u^2+v^2+w^2)/c^2\}^2} \tag{3}$$

The acceleration of a free particle moving with velocity $\mathbf{V}$ (u,v,w), at the position $\mathbf{P}$ (x,y,z), at epoch t is given by

$$\frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi), \tag{4}$$

where $$X = t^2-\frac{\mathbf{P}^2}{c^2}, \qquad Y = 1-\frac{\mathbf{V}^2}{c^2}, \qquad Z = t-\frac{\mathbf{P}\,\mathbf{V}}{c^2}, \tag{5}$$

$$\xi = Z^2/XY \tag{6}$$

The spatial particle-density n_0 near the observer (putting $x = y = z = 0$) is given by

$$n_0 = B/c^3t^3 \tag{7}$$

If m_0 is the mass of a particle (here representing a nebula), the mass-density ρ_0 is given by

$$\rho_0 = n_0 m_0 = m_0 B/c^3t^3 \tag{8}$$

The acceleration of a free particle P near the observer, at rest relative to the observer, is

$$\frac{d\mathbf{V}}{dt} = \frac{\mathbf{P}}{t^2}G(1) \tag{9}$$

The Newtonian acceleration for a particle near the observer is by the inverse square law

$$\frac{d\mathbf{V}}{dt} = -\tfrac{4}{3}\pi\gamma|\mathbf{P}|^3\rho_0\frac{1}{|\mathbf{P}|^2}\frac{\mathbf{P}}{|\mathbf{P}|}, \tag{10}$$

where γ is the Newtonian constant of gravitation. The two values of the acceleration will agree, for $|\mathbf{P}|$ small, if

$$-\tfrac{4}{3}\pi\gamma\rho_0 = \frac{G(1)}{t^2}, \tag{11}$$

i.e., using (8), if

$$\tfrac{4}{3}\pi m_0 B = \frac{c^3 t}{\gamma}[-G(1)]. \tag{12}$$

The number $G(1)$ has been shown to be negative, and by arguments depending on the more detailed statistical world-system $-G(1)$ was shown in certain cases to be unity. The hydrodynamical or simple kinematic system can be considered as the limit of the statistical system for marked nuclear condensation, and without more ado I shall assume that $-G(1)$ is in fact unity for the simple kinematic system.

413. Since B is a constant, if we define m_0 to be constant then we must have $\gamma \propto t$. This means that the Newtonian 'constant' of gravitation, if estimated by this method from the acceleration of a free particle in the space immediately surrounding a not too small group of neighbouring galaxies, should be proportional to the epoch t reckoned from the natural time-zero. It should therefore be apparently increasing at the rate of one part in 2×10^9 per year. This must not be taken to imply that local gravitation, in the solar system for example, possesses a varying γ. Our result relates to the estimate of gravitation made by an observer observing a freely moving particle in motion at small relative velocities in inter-galactic space not too far off.

414. The quantity $\tfrac{4}{3}\pi m_0 B$ has an immediate physical meaning. The density, in the experience of observer O at the origin, increases outwards, ultimately to infinity. But suppose he estimates what may be called the extrapolated mass of the universe by multiplying his local density by the total volume of occupied space as estimated from the maximum distance of nebulae inferred from the velocity-distance proportionality. His local density is $\rho_0 = m_0 B/c^3t^3$, the volume is $\tfrac{4}{3}\pi(ct)^3$. Consequently, the extrapolated homogeneous mass, namely the product of these, is

$$\tfrac{4}{3}\pi m_0 B,$$

i.e.

$$c^3 t/\gamma. \tag{13}$$

For $t = 2\times 10^9$ years $= 0.6\times 10^{17}$ secs., $c = 3\times 10^{10}$ cm. sec.$^{-1}$,

$\gamma = 6\ 66 \times 10^{-8}$, this mass is $2\ 4 \times 10^{55}$ grammes It must not be supposed that this is the mass of the universe, which on our analysis is infinite But the mass c^3t/γ, where t is the observer's present epoch, constantly appears in all relativistic cosmologies, some of which indeed assign it as the actual mass of the universe We shall meet this quantity again in our account of current relativistic cosmology, in which, however, γ is a constant of nature

415. The local density, namely $n_0 m_0$ or $m_0 B/c^3t^3$, is now seen to be

$$\rho_0 = \frac{1}{\frac{4}{3}\pi\gamma t^2}, \tag{14}$$

or 10^{-27} gramme cm $^{-3}$ at the present epoch It is remarkable that this is independent of the arbitrary constant B A density of the order $1/\pi\gamma t^2$ constantly appears as the local density at epoch t in all cosmological treatments, as we shall see

416 A given nebular nucleus in the vicinity of a given group of nebulae is on this model unaccelerated This can only be reconciled with the Newtonian law by introducing a repulsive force proportional to the distance The value of the required constant of proportionality is readily found The equation of motion of the nebula must be

$$0 = -\gamma\frac{M(r)}{r^2} + \tfrac{1}{3}c^2\lambda r, \tag{15}$$

where $\frac{1}{3}c^2\lambda r$ is the repulsive acceleration which is the resultant of all the repulsive actions at a distance arising from the totality of matter present † Now here $M(r) = \frac{4}{3}\pi\rho_0 r^3 = r^3/\gamma t^2$, whence

$$\lambda = \frac{3}{c^2t^2}, \tag{16}$$

independent of r The present numerical value of this is about 10^{-54} cm $^{-2}$ This is approximately the value assigned to the 'cosmical constant' λ in Einstein's original static universe

417. It is evident without calculation that this must come out so For our system, like Einstein's, assigns zero accelerations to the particles or nebular nuclei, though whilst in Einstein's model the

† It is readily shown that a system of repelling point centres, spherically symmetrical about O and each repelling with a force proportional to the distance, gives a resultant repulsion at any point proportional to the distance of the point from O This follows by a simple application of the triangle of forces

nebulae are all at relative rest (in contradiction with the impossibility of describing a standard of absolute rest), in our system the nebulae possess such constant velocities that they determine no preferential standard of rest It must be remembered, however, that in our treatment the 'cosmical constant' (a function of the time) would represent a mere subjective resolution of the accelerations 'zero' into an attraction and an equal and opposite repulsion In Einstein's treatment it was supposed to be a genuine constant of nature, though introduced empirically In our kinematic treatment of gravitational problems no constants of nature are introduced at all Our comparison affords great insight, however, into the manner in which an empirical procedure such as Einstein's can lead to a belief in the objective existence of a quantity which a more deep-going analysis shows to be merely subjective Thus do we dignify with the name 'laws of nature' regularities or relations inserted into the situation by the observer—a view held and forcibly expressed in his Gifford lectures by Eddington, but a view from which his own practice, in relation to the 'cosmical constant', has somewhat deviated If the observer chooses to believe in action at a distance, he will inevitably be led to introduce 'constants' of nature, but these are purely products of his own imagination, at least in this context † Our analysis not only introduces no cosmical constant, but makes it highly improbable that any such constant has any part in the ultimate description of phenomena

418. The foregoing formulae (13), (14), (16) are reproduced here for convenient comparison with the similar formulae of relativistic cosmology We now proceed to state briefly the properties of the simple kinematic system in terms of observations

Possible observations

419. We have seen that the essential instruments of observation are the clock and the theodolite We make an observation on a distant observer's clock, in effect, by measuring the Doppler effect for the lines in the source he accompanies It is therefore sufficient to replace the instruments of observation by a clock, telescope, and spectro-

† The constant γ, of course, serves a useful purpose—it enables us to pass from particle counts to mass measures, just as c enables us to pass from time-measures to length measures, the value of γ fixes the gramme, the value of c the centimetre

scope The following observations can then be made on any world-system

(i) The Doppler shift-coefficient s of any nebula can be measured at *epoch of observation* t_2 by the observer's clock, and so s determined as a function of t_2,† $s \equiv s(t_2)$

(ii) The number Nds of nebulae per unit solid angle with observed Doppler shifts between s and $s+ds$ can be counted at epoch of observation t_2 by the observer's clock This can thus be determined as a function of s at epoch t_2, and by repeating the observations at different epochs it can be determined as a function of the two variables s and t_2, $N \equiv N(s, t_2)$

In addition, (iii) the telescope will yield some information about the state of evolution of the nebula identified by the Doppler shift s This can be interpreted as yielding the local nebular time t' on the nebula defined by s as a function of the epoch of observation t_2, $t' \equiv t'(s, t_2)$

Lastly, (iv) if the observer is provided with a photometer or other brightness-measuring apparatus, he can determine the brightness b of the s-nebula at epoch of observation t_2 as a function of s and t_2, $b \equiv b(s, t_2)$

420. Just as these observations can be actually carried out on the actual universe, so any proposed model of the universe, whatever geometry or laws of nature be adopted, can be taken as predicting the result of these observations Any model proposed for the universe thus predicts at least four functions involving only observed quantities, namely $s(t_2)$, $N(s, t_2)$, $t'(s, t_2)$, and $b(s, t_2)$ Any proposed model can be compared with the actual universe by constructing these four functions mathematically and then comparing them with the results of observation

421. Similarly any two models of the universe can be compared by constructing the four functions mathematically and comparing the functions In such a comparison everything introduced by the observer—geometry, coordinates, assumed laws of nature—is eliminated, and the comparison is one of observations only We compare as it were the actual photographs and spectrograms that would be obtained by exposing large telescopes and spectroscopes to the skies

† It is convenient to continue to use t_2 for the epoch of reception of a light signal, as in Chapter II

It then follows that two proposed models of the universe are identical, and only identical, when these four functions (at least) are identical

422. I propose in due course to construct these functions, or rather the first three of them, for the simple kinematic model we have examined in Part II and for the universes of current relativistic cosmology, and then to compare them This will settle definitely the question whether our models do or do not coincide with the models of current relativity

423. It is surprising that this methodology has not been adopted by previous writers † Investigators have analysed *ad nauseam* the *geometrical* properties of universes on different geometries, with different values of the 'cosmical constant' and differing values of the constants of integration that arise, but have scarcely paused to tabulate the corresponding *observable* properties in detail It is, however, only by comparing observable properties that we can institute any comparisons between the universes of current relativistic cosmology and the kinematic universes constructed on such different principles

424. I shall in fact confine myself to the observational properties (i), (ii), and (iii) Property (iv) is readily discussed, but its discussion requires certain principles of physical optics which are outside the scope of the purely kinematic considerations of the present work Moreover a complete discussion of the luminosity or brightness properties of nebulae requires the making of some assumption as to the variation or constancy of luminosity of a given nebula in its own local time-history These problems will perhaps be discussed in due course in technical papers

Observable functions for the simple kinematic system

425. The simple kinematic system of moving particles is defined by (1) and (2) above The velocity of any particle is constant Consequently its observed Doppler shift is independent of the epoch of observation t_2 The function $s(t_2)$ giving the variation of Doppler

† Since this section was written, the procedure here advocated (and advocated earlier in *Observatory*, 57, 24, 1934) has been carried out by Tolman, *Relativity, Thermodynamics, and Cosmology* (1934), Part IV, and by McCrea, *Zeits fur Astrophys* 9, 290, 1935

shift with epoch of observation is accordingly

$$s(t_2) \equiv \text{constant} \tag{17}$$

This is observable function (i)

426. The coordinate recession velocity V has been shown to be connected with the Doppler shift-ratio s by the relation

$$s = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}} \tag{18}$$

It is worth while, in parenthesis, to derive this again from first principles If a light-signal leaves a particle at epoch t in O's reckoning, when the distance of the particle is r, then by the definitions of epoch and distance it arrives at O at epoch t_2 by O's clock, where

$$t_2 = t+r/c \tag{19}$$

Hence, for two neighbouring signals,

$$\begin{aligned} dt_2 &= dt+dr/c \\ &= dt(1+V/c) \end{aligned}$$

But if t' is the local time on the particle considered, we know that

$$t' = t(1-V^2/c^2)^{\frac{1}{2}}, \qquad dt' = dt(1-V^2/c^2)^{\frac{1}{2}} \tag{20}$$

Hence by the definition of the Doppler shift-ratio s,

$$s = \frac{dt_2}{dt'} = \left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}}$$

Solving, we have
$$\frac{V}{c} = \frac{s^2-1}{s^2+1} \tag{21}$$

Hence for a range dV in V, the corresponding range ds in s is given by

$$\frac{dV}{c} = \frac{4s\,ds}{(s^2+1)^2} \tag{22}$$

Now formula (3) above gives the number of particles with recession velocities between V and dV, inside a solid angle $d\omega$, as

$$\frac{BV^2\,dV d\omega}{c^3(1-V^2/c^2)^2} \tag{23}$$

Introducing for V and dV in terms of s and ds from (21) and (22) in (23), the desired number is

$$B\frac{(s^2-1)^2}{4s^3}\,dsd\omega \tag{24}$$

This is the observable function (ii)

427. The distant local time t' is connected with the time of observation t_2 by the formulae (19) and (20) Putting $r = Vt$ we have

$$t_2 = t(1+V/c) = t'\left(\frac{1+V/c}{1-V/c}\right)^{\frac{1}{2}} = st'$$

or
$$t' = t_2/s \tag{25}$$

This is the observable function (iii)

428. Our three functions connecting observables are accordingly

$$s = \text{const (indep of } t_2), \tag{i}$$

$$N\,d\omega ds = B\frac{(s^2-1)^2}{4s^3}dsd\omega, \tag{ii}$$

$$t' = t_2/s \tag{iii}$$

In these t_2 is the epoch at the observer at which the observation is made, and s is the Doppler shift-coefficient observed at this epoch Formula (i) gives the march of s for any arbitrary nebula, (ii) gives the density of particles on a photo-spectrogram in terms of s and epoch t_2, and (iii) gives the local time t', on the nebula identified by s, at epoch of observation t_2

429. In Chapter XVII we shall obtain the corresponding formulae, in terms of the same observable variables s and t_2, for the universes of general relativity In order to obtain these formulae in as simple a way as possible, we turn aside in Chapter XVI to consider how the cosmological problem would be solved by a Newtonian analyst using Newtonian time and Euclidian space for all observers but assuming these times and spaces to be connected not by the correct formulae of relativity but by the simple (incorrect) formulae of Newtonian relativity

XVI

NEWTONIAN AND RELATIVISTIC COSMOLOGY

430 Great insight into the cosmological problem is obtained by constructing a solution by the methods of classical mechanics (Newtonian dynamics and gravitation) and Newtonian relativity This will be found to convey an easily intelligible idea of the *dynamical* nature of the phenomenon of the expanding universe Further it will be found greatly to facilitate the mathematical discussion of the dynamical relativistic solutions, by suggesting appropriate mathematical procedures and by attaching physical meanings to certain of the symbols occurring in the dynamical relativistic solutions whilst indicating at the same time what symbols are devoid of physical content

431 By a Newtonian solution of the cosmological problem we mean of course a distribution of matter in motion satisfying the cosmological principle and obeying the laws of Newtonian dynamics and gravitation in accordance with Newtonian relativity In Newtonian mechanics, once a system is *given*, its future motion is determinate The object of the investigation is to find such a system of moving particles such that for any pair of particles A, B both $A \equiv B$ and $A ≣ B$, or, in words, such that the description of the system from any particle A of the system as observing head-quarters is identical with the description of the system from any other particle B of the system This is of course the hydrodynamical case, in which the members of *every* pair of particles are equivalent, the more refined statistical case, in which the system contains particles besides the equivalent pairs, I shall not here consider, though there is no difficulty in constructing one The addition of a *dynamics* to the problem of course enormously simplifies the mathematics as compared with the solution by purely kinematic methods, where we have no empirical dynamics to guide us

432 There are various other advantages in constructing solutions of the cosmological problem by the methods that would have been used by an analyst of the Newtonian period First, it was until recently generally supposed that a solution of the cosmological problem on Newtonian principles was impossible This was the view of Einstein

when he wrote † 'If we ponder over the question as to how the universe, considered as a whole, is to be regarded, the first answer that suggests itself is surely this As regards space (and time) the universe is surely infinite There are stars everywhere, so that the density of matter, although very variable in detail, is nevertheless on the average everywhere the same In other words However far we might travel through space we should find everywhere an attenuated swarm of fixed stars of approximately the same kind and density — This view is not in harmony with the theory of Newton The latter theory rather requires that the universe should have a kind of centre in which the density of stars is a maximum, and that as we proceed outwards from this centre the group-density of the stars should diminish, until finally, at great distances, it is succeeded by an infinite region of emptiness The stellar universe ought to be a finite island in the infinite ocean of space —This conception is of itself not very satisfactory It is still less satisfactory because it leads to the result that the light emitted by the stars and also individual stars of the stellar system are perpetually passing out into infinite space, never to return, and without ever again coming into interaction with other objects of nature Such a finite material universe would be destined to become gradually but systematically impoverished ' Einstein appended a proof claiming to show that the intensity of the gravitational field at the surface of any sphere in a homogeneous infinite distribution of matter would ultimately become infinite with increasing radius of the sphere, 'which is impossible' Elsewhere‡ Einstein remarked 'Diese Schwierigkeiten lassen sich auf dem Boden der Newtonschen Theorie wohl kaum uberwinden '

The above considerations are, however, invalid In the first place, the notion of the 'intensity' of a gravitational field is a pure concept Einstein's mathematical proof involved the concept of lines of force But the point is not whether the conceptual field, calculated with hypothetical boundary conditions, in a given frame of reference, ever becomes infinite, but whether the accelerations, actually undergone by the particles present and capable of observation by other particle-observers, become infinite in the experience of these particle-observers, reckoned in the frame moving with the particle-observer considered

† *The Theory of Relativity* (English trans , Methuen), 4th edition, chapter 30, p 105 (1921)

‡ *Sitz Preuss Akad Wiss* , 1917, S 150

These have to be measured by specified particle-observers before they have a meaning, and it can be shown that Newtonian systems can be constructed in which the relative accelerations are small near the observer and everywhere finite in his experience In the second place, there is no need to posit, in Newtonian mechanics, a unique centre for the universe, with density everywhere decreasing outwards, Newtonian mechanics by no means implies the 'island universe' view This will be shown by actual construction of systems which are both homogeneous (in the Newtonian sense) and self-consistent, i e which obey everywhere the laws of Newtonian mechanics In the third place Einstein's argument claiming to deny the likelihood of particles and light passing away from one another for ever and claiming to demand perpetual reinteraction is of a metaphysical, *a priori* character, and we have no justification for imposing this demand (Actually our statistical kinematic system satisfies Einstein's demand, in part, but it does so as an *a posteriori* consequence of our application of the principle of relativity)

Einstein's considerations were put forward before the phenomenon of the expansion of the universe was generally recognized They show how dangerous it is to attach weight to *a priori* general arguments The existence of *motion* as well as *matter* in the universe is an essential element which Einstein at that time left out It is indeed impossible to construct a *static* world-system satisfying Newtonian dynamics without modifications of Newton's law of gravitation, such as that which was suggested by Seeliger, or that which was suggested by Einstein which was equivalent to the introduction of the 'λ-term' in his field equations With the important proviso that motion is left out, Einstein's argument as to the impossibility of a homogeneous Newtonian universe is valid But as soon as we permit *motion* in the smoothed-out distribution of matter, systems built up upon Newtonian mechanics and on the strict Newtonian law of gravitation become possible, and are fully valid within the domain of Newtonian time and space, that is to say, within the domain of Newtonian relativity They are valid subject to the conception of world-wide time-keeping, the same for all observers Of course, as soon as we inquire by what observations observers can determine this time, we find that observers who are equivalent in the Newtonian sense fail to find themselves equivalent when they employ identical procedures for calculating coordinates (in particular, epochs of distant events) out of observa-

tions, and so the Newtonian concept of a universal time has to be abandoned Instead we have to reconsider the observations which the equivalent observers, using the same rules for combining observations, make on one another, with the result that the Newtonian scheme requires modification The resulting modifications have been the theme of Parts I, II, and III of this work

433. But, secondly, amongst the advantages of a Newtonian dynamical treatment, it will be shown that the modifications required only affect phenomena beyond the immediate neighbourhood of the observer, i e in practice beyond the present range of astronomical instruments, say 200 million light-years' distance from the observer If we call the phenomena occurring within this range of observation 'local phenomena', we may say that the local phenomena predicted by the Newtonian scheme are indistinguishable from the local phenomena predicted by the schemes of current relativistic cosmology Small modifications are indeed introduced if we add to that generalization of the Newtonian scheme which is obtained by taking $G_{\mu\nu}-\frac{1}{2}g_{\mu\nu}G+\kappa T_{\mu\nu}=0$ as the 'field equations' a term $\lambda g_{\mu\nu}$ on the left-hand side But this merely corresponds to modifying the Newtonian inverse square law by the addition of a repulsive acceleration $\frac{1}{3}\lambda r$ $(\lambda > 0)$, and with this modification the Newtonian and the general relativity phenomena are again locally identical Thus Newtonian *theory* (as distinguished from the specific Newtonian law of gravitation) is fully competent to predict a set of locally observable phenomena, whether the local phenomena *actually* observed agree with these predictions is a matter for investigation Thus the astronomer in his discussions of his own observations can justifiably compare them, as far as is yet required by his range of observations, with the theoretical predictions of Newtonian theory, without recourse to general relativity, or 'curved' or 'expanding' space The elaborate calculus employed in the methods of general relativity conceals beneath a cloud of symbols the simple phenomena actually predicted, and robs us of any physical insight into these phenomena For local phenomena, Newtonian time and space are sufficient for the astronomer's purposes until he is driven to consider the kinematical bases of dynamics

434 This remarkable equivalence of local phenomena on Newtonian and relativity dynamics would alone justify a close consideration of

world-systems possible within Newtonian cosmology But, thirdly, a still more remarkable result† will be established It will be shown that the Newtonian differential equations expressing the motion of the particles and their distribution, considered as differential equations with regard to Newtonian time t as independent variable, are identical in form with the relativistic differential equations expressing the motion and distribution, considered as differential equations with respect to 'cosmic time' t' as independent variable, t', it will be remembered, is the time of an event as observed by the moving observer at whom it occurs Newtonian 'distance' r is the same function of Newtonian time t as the conventional relativity coordinate r is of cosmic time t', or, apart from an arbitrary multiplying constant, the same function as the conventional 'radius of curvature of space' R The inner meaning of this remarkable identity of form remains for future investigation It indicates some deep-lying parallelism between the forms of thought employed in Newtonian and in relativistic dynamics, when these are applied to the construction of a cosmology Whether local, isolated gravitational situations (such as the Keplerian problem) can be thrown into formally identical symbolisms on the two schemes is not yet known, the symbolisms at present customarily employed exhibit them, of course, differently In the meantime the parallelism allows us to attach physical meanings to the symbols occurring in the relativistic equations, which symbols have been assigned hitherto a purely mathematical content The interpretation of the relativistic equations is thus greatly facilitated The interpretations of the identical Newtonian and relativistic symbolisms are of course different, the imperfect code of interpretation of the Newtonian symbols into observations has to be replaced by the self-consistent relativistic interpretation

435. When we have thus constructed Newtonian world-systems, and effected a complete correspondence between them and the current relativistic solutions, and when we have correctly interpreted the relativistic formalism in terms of observations, we can proceed to compare the current relativistic world-pictures with the kinematic world-pictures We now go to analytical details

† This general result is due to W H McCrea and the author, working in collaboration, after particular cases of it had been discovered by the author alone

The Newtonian 'parabolic' universe†

436 If a system of particles in motion of hydrodynamical character satisfies the cosmological principle on Newtonian mechanics, then it must be homogeneous in density For each of the equivalent particle-observers uses the same world-wide time and space, if t and t' are the epochs of an event to two Newtonian observers O and O', then $t' = t$, and $\rho(A, t)$ has an unambiguous meaning, and for two such observers to possess identical descriptions of the system in terms of coordinates measured from themselves, we must have $\rho(A, t) = \rho(B, t)$ independent of the positions of A and B We have already seen (Chapter III) that absolute homogeneity has an unambiguous meaning in Newtonian time But the density is not necessarily constant in time We have already shown (Chapter IV) that in a Newtonian system satisfying the cosmological principle, the motion of every particle must be radially outward (or inward) from every observer, and we have seen that the velocity v must be of the form $rF(t)$, where r is the Newtonian distance Our object is to determine the forms of $F(t)$ and of $\rho(t)$ implied by Newtonian mechanics

437. To be precise, let v be the particle-velocity, radial in direction, at distance r from an observer O situated on one of the particles of the system Let $M(r)$ be the mass contained in the sphere of radius r, centre O We gain insight by considering first a particular case Consider the case in which, on Newtonian mechanics, the particle at r has the parabolic velocity of escape from the mass contained within the sphere of radius r If the gravitational effect of the material outside the sphere of radius r can be ignored, then if the particle once possesses the parabolic velocity of escape, it does so always For the particle moves with the boundary of this sphere, and the mass inside the sphere is constant, consequently the motion, on the assumption made, is simply that of a particle just escaping from a fixed mass We shall have to verify that the assumption is in fact legitimate, i e that for all observers the material outside their respective spheres, for any r, at any t, can be self-consistently ignored

438. The parabolic velocity of escape v is defined by the equation

$$\tfrac{1}{2}v^2 = \frac{\gamma M(r)}{r}, \tag{1}$$

† *Quart J of Math* (Oxford), **5**, 64, 1934

where γ is the Newtonian constant of gravitation This definition involves no appeal to the notion of a gravitational potential, which is here inapplicable since we can never get to a great distance from all attracting matter It is merely the integral of the Newtonian equation of motion

$$\frac{Dv}{Dt} = -\frac{\gamma M(r)}{r^2}, \tag{2}$$

obtained by multiplying by v or Dr/Dt, integrating, and setting the constant of integration equal to zero Since $M(r) = \frac{4}{3}\pi r^3 \rho$, equation (1) gives

$$v^2 = \tfrac{8}{3}\pi\gamma r^2 \rho \tag{3}$$

The motion must be such that the hydrodynamical equation of continuity—the èquation expressing the conservation of mass—must be satisfied This equation runs, in polar coordinates and Eulerian notation,

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \rho v) = 0 \tag{4}$$

Here, v is a function of r and t given by (3) Inserting (3) in (4), since $\rho \equiv \rho(t)$ is a function of t only, we have

$$\rho^{-\frac{3}{2}}\frac{d\rho}{dt} + 3\left(\frac{8\pi\gamma}{3}\right)^{\frac{1}{2}} = 0$$

Integrating, we have

$$-2\rho^{-\frac{1}{2}} + (24\pi\gamma)^{\frac{1}{2}} t = \text{const} ,$$

or, choosing a suitable origin of t,

$$-2\rho^{-\frac{1}{2}} + (24\pi\gamma)^{\frac{1}{2}} t = 0$$

This gives

$$\rho = \frac{1}{6\pi\gamma t^2}, \tag{5}$$

and (3) then gives†

$$v = \frac{2}{3}\frac{r}{t}$$

439 We now verify that this is a solution of the dynamical problem The acceleration of the particle is

$$\frac{Dv}{Dt} = \frac{D}{Dt}\left(\frac{2}{3}\frac{r}{t}\right) = \frac{2}{3}\left(\frac{v}{t} - \frac{r}{t^2}\right) = -\frac{2}{9}\frac{r}{t^2},$$

and this is precisely the Newtonian acceleration required by (2), since

$$-\gamma\frac{M(r)}{r^2} = -\gamma\frac{\frac{4}{3}\pi r^3}{r^2}\frac{1}{6\pi\gamma t^2} = -\frac{2}{9}\frac{r}{t^2}$$

† The minus sign is also permissible, and gives a contracting universe The choice of the positive sign can be made on the basis of observation, or, more fundamentally, on the basis of the arguments of Chapter IV, which established the *a priori* improbability of a contracting universe in certain cases

440. Now consider any second observer O' attached to some other moving particle of the system. Let his distance from O at time t be the vector $\mathbf{R}$. Then his velocity is $\frac{2}{3}\mathbf{R}/t$, and his acceleration $-\frac{2}{9}\mathbf{R}/t^2$, as reckoned by O. Consequently the distance $\mathbf{r}'$, velocity $\mathbf{v}'$, and acceleration $\mathbf{g}'$, relative to himself, which he assigns to the distant particle whose distance, velocity, and acceleration are respectively $\mathbf{r}$, $\mathbf{v}$, and $\mathbf{g}$ as reckoned by O, are given by

$$\mathbf{r}' = \mathbf{r} - \mathbf{R},$$

$$\mathbf{v}' = \mathbf{v} - \frac{2}{3}\frac{\mathbf{R}}{t},$$

$$\mathbf{g}' = \mathbf{g} + \frac{2}{9}\frac{\mathbf{R}}{t^2},$$

by the principles of Newtonian relativity. Hence

$$\mathbf{v}' = \frac{2}{3}\left(\frac{\mathbf{r}}{t} - \frac{\mathbf{R}}{t}\right) = \frac{2}{3}\frac{\mathbf{r}'}{t},$$

$$\mathbf{g}' = -\frac{2}{9}\left(\frac{\mathbf{r}}{t^2} - \frac{\mathbf{R}}{t^2}\right) = -\frac{2}{9}\frac{\mathbf{r}'}{t^2},$$

and the velocity and accelerations are thus of the same form in O''s coordinates as they were in those of O. Consequently O''s measures of the motion of the particle considered satisfy the same dynamical conditions with respect to O' as those of O did with respect to O. Thus O's assumption that the matter outside the sphere of radius r, centre O, can be ignored implies the same assumption for O', and so is self-consistent. Each particle in the Newtonian system can equally be regarded as the 'centre' of the whole system, it behaves as if distributed with spherical symmetry round every particle.

441. It will be noticed that the velocity law $v = \frac{2}{3}r/t$ and density-law $\rho = 1/6\pi\gamma t^2$ are of the same general form as the velocity-law $v = r/t$ and the *local* density law $\rho_0 = 1/\frac{4}{3}\pi\gamma t^2$ on the simple kinematic system, with this difference, that the simple kinematic system is only capable of local Newtonian description if we take $\gamma \propto t$. The numerical values of the local density calculated via the observed recession-law is, however, much the same as for the simple kinematic system. For if we write as the velocity-law $v = \alpha r$, where α is the present observed value of the recession constant, then the parabolic Newtonian system gives $t = \frac{2}{3}\alpha^{-1}$, $\rho = \alpha^2/\frac{8}{3}\pi\gamma$, whilst the simple

kinematic system gives $t = \alpha^{-1}$, $\rho = \alpha^2/\frac{4}{3}\pi\gamma$ The estimate of the 'age' is multiplied by $\frac{2}{3}$, the estimate of the density by $\frac{1}{2}$

442. We notice that the velocity law for the Newtonian parabolic universe, namely $v = \frac{2}{3}r/t$, is of the form $v = rF(t)$, as it must be on any cosmology satisfying the equation of continuity with a suitable meaning for t

Comparison with the relativistic universe of Einstein and de Sitter

443. The particular Newtonian universe thus constructed is effectively defined by the two equations

$$\tfrac{1}{2}v^2 = \frac{\gamma M(r)}{r}, \qquad \frac{Dv}{Dt} = -\frac{\gamma M(r)}{r^2}, \tag{6}$$

with
$$M(r) = \tfrac{4}{3}\pi\rho r^3, \tag{7}$$

for equations (6) themselves imply the equation of continuity This is seen by differentiating the first of equations (6), when we get

$$v\frac{Dv}{Dt} = -\frac{\gamma M(r)}{r^2}v + \frac{\gamma}{r}\frac{D}{Dt}M(r),$$

and use of the second of (6) yields at once

$$\frac{D}{Dt}M(r) = 0$$

Integration of $Dr/Dt = \frac{2}{3}r/t$ gives at once for the motion

$$r = \text{const } t^{\frac{2}{3}} \tag{8}$$

Now transform (6) by writing

$$r = fR(t) \qquad (R(t) \propto t^{\frac{2}{3}}), \tag{9}$$

where the value of f, corresponding to the arbitrary constant in (8), specifies the particle under consideration Then

$$v = f\frac{dR}{dt}, \qquad \frac{Dv}{Dt} = f\frac{d^2R}{dt^2}, \qquad M(r) = \tfrac{4}{3}\pi\rho f^3R^3$$

Introducing these in (6) we have

$$\left(\frac{1}{R}\frac{dR}{dt}\right)^2 = \frac{8\pi\gamma}{3}\rho, \tag{10}$$

and
$$\frac{2}{R}\frac{d^2R}{dt^2} = -\frac{8\pi\gamma}{3}\rho$$

$$= -\left(\frac{1}{R}\frac{dR}{dt}\right)^2, \tag{11}$$

equations independent of f, and so expressing properties true for all particles of the system

444. Next introduce 'Einstein's constant' κ defined by

$$\kappa = 8\pi\gamma/c^2,$$

and transform the time-variable by writing

$$c\,dt = d\tau$$

Then (10) and (11) become

$$\left(\frac{1}{R}\frac{dR}{d\tau}\right)^2 = \tfrac{1}{3}\kappa\rho, \tag{12}$$

$$\frac{2}{R}\frac{d^2R}{d\tau^2}+\left(\frac{1}{R}\frac{dR}{d\tau}\right)^2 = 0 \tag{13}$$

Equations (12) and (13) are identical with the equations for an expanding universe of zero curvature with pressure $p = 0$ and cosmical constant $\lambda = 0$ defined by the metric

$$ds^2 = d\tau^2-[R(\tau)]^2(dx^2+dy^2+dz^2) \tag{14}$$

This is the well-known model first considered by Einstein and de Sitter †

445. Conversely, from the relativistic equations (12) and (13) found by Einstein and de Sitter we can infer equations identical *in form* with the parabolic Newtonian equations (6) and (7) For the 'coordinate distance' r of a particle from the origin in the metric (14) is given by

$$r = fR,$$

where f is a constant for the particle, depending on the particle chosen Then

$$\frac{1}{r}\frac{dr}{d\tau} = \frac{1}{R}\frac{dR}{d\tau}$$

Introducing these in (12) and (13), returning to t and substituting for κ we see that c cancels and we are left with

$$\left(\frac{1}{r}\frac{dr}{dt}\right)^2 = \frac{8\pi\gamma}{3}\rho, \tag{15}$$

$$\left(\frac{2}{r}\frac{d^2r}{dt^2}\right)+\left(\frac{1}{r}\frac{dr}{dt}\right)^2 = 0 \tag{16}$$

If we define $m(r)$ by $$m(r) = \tfrac{4}{3}\pi r^3\rho, \tag{17}$$

† *Proc Nat Acad Sci*, **18**, 213, 1932, *Univ of California Pub Math*, **2**, 161, 1933

then by (15)

$$\frac{D}{Dt}m(r) = \frac{D}{Dt}(\tfrac{4}{3}\pi r^2\rho\ r)$$
$$= \frac{1}{2\gamma}\frac{D}{Dt}\left\{r\left(\frac{dr}{dt}\right)^2\right\},$$

which by (16) is zero Hence $m(r)$ is constant following the motion, and (15) and (16) then give

$$\tfrac{1}{2}v^2 = \frac{\gamma m(r)}{r}, \qquad \frac{Dv}{Dt} = -\frac{\gamma m(r)}{r^2}, \tag{18}$$

which are identical in form with the Newtonian equations It then follows, exactly as in the Newtonian case, that for the system of Einstein and de Sitter the density ρ is given in terms of cosmic time t by the formula $\rho = 1/6\pi\gamma t^2$, and that the coordinate velocity (rate of change of coordinate distance with respect to cosmic time) is given by $v = \frac{2}{3}r/t$

446. The exact observational interpretation of the Einstein-de Sitter system will be found in the next chapter In the meantime we notice, first, that since cosmic time t (the time kept by each particle-observer for events at himself) coincides locally with Newtonian time t, the identity of the formal relations shows that all local phenomena are identical in the parabolic Newtonian system and in the Einstein-de Sitter system It follows that an analyst of Newton's period, unacquainted with general relativity, who had formulated on Newtonian principles the parabolic Newtonian universe would have been led to predict the same local phenomena as an investigator adopting Einstein and de Sitter's procedure He would have been led to the same predicted relation between the local rate of expansion and the local density, and to the same estimate of the local age reckoned from $t = 0$

447. Secondly, there is a perfect and complete correspondence between the parabolic Newtonian and the Einstein-de Sitter relativistic universes, valid not only locally but for all values of the coordinates A single example of this correspondence will now be given

448. In the parabolic Newtonian universe, each particle, since it is endowed with the parabolic velocity of escape, comes ultimately to rest relative to the observer as time advances, as $t \to \infty$, $v \to 0$ This is clear since $v = \frac{2}{3}r/t = \frac{2}{3}t^{-\frac{1}{3}} \times \text{const}$

In the corresponding relativistic universe velocities have only a conventional meaning, and we accordingly go straight to observed Doppler effects Let a light-signal leave a particle at local or cosmic time t', and arrive at the observer (at the origin) at epoch of observation t_2 We shall now use r to denote not the coordinate *distance* but the radial coordinate measure, so that $dr^2 = dx^2+dy^2+dz^2$ Then in the space-time of metric

$$ds^2 = c^2\,dt^2-[R(t)]^2\,dr^2, \tag{19}$$

along a light-path $ds = 0$ we have

$$dr = \frac{c\,dt}{R(t)}.$$

Hence if r refers to the particle from which the light-signal arrives, we have

$$r = \int\limits_{t'}^{t_2} \frac{c\,dt}{R(t)} \tag{20}$$

Hence for two neighbouring light-signals from the same particle, of fixed radial coordinate r,

$$0 = \frac{dt_2}{R(t_2)}-\frac{dt'}{R(t')}$$

Hence the Doppler shift coefficient s is given by

$$s = \frac{dt_2}{dt'} = \frac{R(t_2)}{R(t')} \tag{21}$$

Putting $$R(t) = bt^{\frac{2}{3}},$$

where b is arbitrary, (20) and (21) give

$$r/c = 3(t_2^{\frac{1}{3}}-t'^{\frac{1}{3}})/b, \tag{22}$$

$$s = t_2^{\frac{2}{3}}/t'^{\frac{2}{3}} \tag{23}$$

Eliminating t' between these,

$$s = \frac{t_2^{\frac{2}{3}}}{\{t_2^{\frac{1}{3}}-(br/3c)\}^2} \tag{24}$$

For a fixed particle (or nebula) r is constant, and as the epoch of observation t_2 tends to infinity, $s \to 1$

Thus in the Einstein-de Sitter universe, as the epoch of observation advances, the Doppler *displacement* for any given particle tends to zero. The particle thus appears to come ultimately to rest relative to the observer, as in the corresponding Newtonian universe In the relativistic universe the Doppler shift steadily decreases as the

epoch of observation advances, and this will be interpreted by the observer as a steady deceleration

Thus in both the Newtonian parabolic universe and in its relativistic analogue, a state of ultimate rest, or stagnation, supervenes for any given set of objects observed by a given observer This is true, however extensive the observed set of particles The observer thus picks out ultimately an absolute standard of rest This is sufficient of itself to rule out the parabolic relativistic universe as a possible representation of the universe, for though the parabolic relativistic universe is constructed rigorously in accordance with Einstein's embodiment of the principle of relativity, this system does not satisfy the wider demand that it should pick out no unique standard of rest Thus the parabolic relativistic model of Einstein and de Sitter fails to achieve the construction of a universe free from an absolute standard of rest, a demand which is, however, satisfied by the simple kinematic system The circumstance that both systems use a flat space must not mislead the reader—it is a merely superficial resemblance, the 'space' of the simple kinematic system is something very different from any spatial cross-section '$t =$ const ' of the space-time defined by the metric (14) Events are mapped in totally different ways in the two schemes, and the trajectories of receding particles, i e observationally s as a function of t_2, are totally different in the two cases In the next chapter we shall find a much weightier reason for rejecting Einstein and de Sitter's system as a possible representation of the universe in accordance with experience

Hyperbolic and elliptic Newtonian universes†

449. We proceed to obtain the most general Newtonian universes As before, let v denote the Newtonian velocity of a particle at distance r from a particle-observer O of the system at time t, and let $M(r)$ be the mass inside the sphere of radius r

We repeat for the convenience of the reader some of the analysis of Chapter IV In the equation of continuity

$$\frac{\partial\rho}{\partial t}+\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\rho v) = 0, \tag{25}$$

ρ is a function of t only, independent of r Put

$$\frac{1}{\rho}\frac{d\rho}{dt} = -3F(t) \tag{26}$$

† W H McCrea and E A Milne, *Quart J of Math* (Oxford), 5, 73, 1934

Then (25) becomes
$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 v) = 3F(t),$$

whence, integrating,
$$v = rF(t)+\frac{G(t)}{r^2}, \tag{27}$$

where $G(t)$ is an arbitrary function. Conditions of continuity at $r = 0$ require $G(t) \equiv 0$, as we saw in Chapter IV. Here we shall infer the same result by another argument.

The Newtonian equation of motion is
$$\frac{Dv}{Dt} = -\frac{\gamma M(r)}{r^2}, \tag{28}$$

in forming this equation the observer O has tentatively made the same assumption as before, namely that he can regard himself as the centre of spherical symmetry of the whole system. Equation (28) may be written in Eulerian notation, putting $M(r) = \frac{4}{3}\pi r^3\rho$,
$$\frac{\partial v}{\partial t}+v\frac{\partial v}{\partial r} = -\tfrac{4}{3}\pi\gamma r\rho \tag{29}$$

Hence
$$\frac{1}{r}\left[\frac{\partial v}{\partial t}+v\frac{\partial v}{\partial r}\right]$$
must be a function of t only. Inserting for v from (27) and carrying out the necessary partial differentiations we find at once that $G(t) \equiv 0$. Hence
$$v = rF(t) \tag{30}$$

Inserting this in (29), we have
$$F'(t)+[F(t)]^2 = -\tfrac{4}{3}\pi\gamma\rho \tag{31}$$

Equation (30) may be written in the form
$$\frac{1}{r}\frac{dr}{dt} = F(t), \tag{32}$$

where here without ambiguity d/dt denotes differentiation following the motion, i.e. differentiation of r as a function of t along the trajectory of the particle. The integral of (32) is of the form
$$r = fR(t), \tag{33}$$

where f is an arbitrary constant defining the particle considered and $R(t)$ is a universal function of t, independent of the particle considered. (For by (31) $F(t)$ is a universal function of t.) The function $R(t)$, by (26) and (32), satisfies the relations
$$\frac{1}{R}\frac{dR}{dt} = F(t) = -\frac{1}{3}\frac{1}{\rho}\frac{d\rho}{dt} \tag{34}$$

Integrating the equality of the first and last of these, we have

$$\rho = \frac{b}{R^3}, \tag{35}$$

where b is constant Introducing (35) and (34) for ρ and $F(t)$ in (31) we find

$$\frac{1}{R}\frac{d^2R}{dt^2} = -\frac{\frac{4}{3}\pi\gamma b}{R^3},$$

of which the first integral is

$$\left(\frac{dR}{dt}\right)^2 = \frac{\frac{8}{3}\pi\gamma b}{R} + K, \tag{36}$$

where K is constant It follows from (34) that $F(t)$ is given by†

$$F(t) = \left[\frac{\frac{8}{3}\pi\gamma b}{R^3} + \frac{K}{R^2}\right]^{\frac{1}{2}}$$

$$= [\tfrac{8}{3}\pi\gamma\rho + A\rho^{\frac{2}{3}}]^{\frac{1}{2}}, \tag{37}$$

where $$K = Ab^{\frac{2}{3}} \tag{38}$$

Accordingly, by (30), $$v = r[\tfrac{8}{3}\pi\gamma\rho + A\rho^{\frac{2}{3}}]^{\frac{1}{2}} \tag{39}$$

This is simply the most general Newtonian integral of the motion, for since $M(r) = \frac{4}{3}\pi\rho r^3$, it may be written

$$\tfrac{1}{2}v^2 = \frac{\gamma M(r)}{r} + \tfrac{1}{2}A\left(\frac{3}{4\pi}\right)^{\frac{2}{3}}[M(r)]^{\frac{2}{3}}, \tag{40}$$

and $M(r)$ is constant following the motion, hence (39) differentiates into (28)

450. It follows that the given particle possesses the elliptic, parabolic, or hyperbolic velocity according as $A \lesseqgtr 0$ By (36) the constant A is the same for all particles Hence if one particle possesses an elliptic velocity, so do they all, and similarly for the parabolic and hyperbolic cases By (39), v obeys a velocity-distance proportionality, and (39) gives explicitly the density in terms of the ratio of velocity to distance But the density is not uniquely determined—it depends on the constant A Thus the solution is in the form of a continuum of universes, fixed by the values assigned to A

451. It remains to find the density as a function of t This is equivalent to the determination of R as a function of t, which is obtained by integrating (36) It gives

$$t = \int\limits_0^R \frac{R^{\frac{1}{2}}\,dR}{[\frac{8}{3}\pi\gamma b + KR]^{\frac{1}{2}}} \tag{41}$$

† We again choose the positive sign

It is convenient to simplify this by writing

$$\rho = \frac{b}{R^3} = \frac{1}{\theta^3} \tag{42}$$

Then

$$t = \int\limits_0^\theta \frac{\theta^{\frac{1}{2}}\,d\theta}{(\frac{8}{3}\pi\gamma + A\theta)^{\frac{1}{2}}} \tag{43}$$

452. It will be observed that b occurs only in combination with R, and that it disappears from the formulae (39) and (43) expressing physical relationships In fact b and R are purely conventional magnitudes devoid of any special physical significance, but b can be chosen so that R is a length The Newtonian universes depend on the single physical parameter A

When $A = 0$, (43) gives

$$t = \tfrac{2}{3}\theta^{\frac{3}{2}}(\tfrac{8}{3}\pi\gamma)^{-\frac{1}{2}},$$

or

$$\rho = \frac{1}{\theta^3} = \frac{1}{6\pi\gamma t^2},$$

whilst (39) gives then

$$v = \frac{2}{3}\frac{r}{t},$$

in agreement with our special treatment for the parabolic case Also $R(t) = b^{\frac{1}{3}}\theta = (6\pi\gamma b)^{\frac{1}{3}}t^{\frac{2}{3}}$

Comparison with the universes of current relativistic cosmology

453 The equation preceding (36) may be written

$$\frac{2}{R}\frac{d^2R}{c^2dt^2} = -\tfrac{1}{3}\kappa\rho, \tag{44}$$

where κ is 'Einstein's constant', given by

$$\kappa = 8\pi\gamma/c^2$$

Also (36) may be written

$$\left(\frac{1}{R}\frac{dR}{c\,dt}\right)^2 + \frac{k}{R^2} = \tfrac{1}{3}\kappa\rho, \tag{45}$$

where

$$k = -\frac{K}{c^2} = -\frac{Ab^{\frac{2}{3}}}{c^2} \tag{46}$$

Adding (44) and (45) we have

$$\frac{2}{R}\frac{d^2R}{c^2\,dt^2} + \left(\frac{1}{R}\frac{dR}{c\,dt}\right)^2 + \frac{k}{R^2} = 0 \tag{47}$$

Equations (45) and (47) are formally identical with the equations of current relativistic cosmology for an expanding universe of 'radius' R, space-curvature k/R^2, with cosmical constant $\lambda = 0$ and pressure $p = 0$ Such a universe is defined by the metric

$$ds^2 = c^2\,dt^2 - R^2\,d\sigma^2, \tag{48}$$

where
$$d\sigma^2 = \frac{dx^2+dy^2+dz^2}{[1+\frac{1}{4}k(x^2+y^2+z^2)]^2}, \tag{49}$$

and equations (45), (47) are consequences of the field-equations

$$G_{\mu\nu}-(\tfrac{1}{2}G-\lambda)g_{\mu\nu} = -\kappa T_{\mu\nu} \tag{50}$$

applied to this metric † In the relativistic equations (45) and (47) t denotes cosmic time, the time kept by the moving particle-observer experiencing the event.

454. It follows that the Newtonian distance r of a particle from the observer is the same function of Newtonian time t as the relativistic function R is of cosmic time t (The relativistic function R is proportional to the coordinate distance of a particle of fixed coordinates) There is a complete correspondence between the Newtonian universes and the relativistic universes in the sense that a Newtonian universe of elliptic velocity ($A < 0$) corresponds to a relativistic universe of positive space-curvature ($k > 0$), a Newtonian universe of hyperbolic velocity ($A > 0$) to a relativistic universe of negative space-curvature ($k < 0$) It follows also that the local properties of the relativistic universes are identical with the local properties of the corresponding Newtonian universes, the *local* world-picture given by observation is the same in the two schemes

Physical interpretation of the 'curvature of space' in relativistic cosmology

455. The correspondence can be put in a still more striking way Consider a Newtonian universe of given A Put v_0 equal to the parabolic velocity in the same universe at distance r Then by (39) or (40), (35) and (46), we have

$$\frac{v^2-v_0^2}{r^2} = A\rho^{\frac{2}{3}} = \frac{Ab^{\frac{2}{3}}}{R^2} = -c^2\frac{k}{R^2}$$

† See, e g , Robertson, 'Relativistic Cosmology', *Reviews of Modern Physics*, **5**, 66, Jan 1933, or de Sitter, *Univ of California Pub Math* , **2**, 161, 1933

Put $1/\mathfrak{R}^2$ equal to the curvature k/R^2 of the corresponding relativistic universe Then the last relation may be written

$$\frac{v^2-v_0^2}{c^2} = -\frac{r^2}{\mathfrak{R}^2} \tag{51}$$

This simple equality contains the full physical content of the *curvature* of a relativistic universe, $\mathfrak{R}^2$ is positive or negative according as $v^2 \lesseqgtr v_0^2$, and the conventional value of $\mathfrak{R}^2$ bears to the square of the Newtonian distance r the same ratio as the conventional value c^2 of the (velocity of light)^2 bears to the deficiency of v^2 below v_0^2 The space-curvature of a relativistic universe is merely a mathematical symbol representing the relation of the actual velocity of expansion to the parabolic velocity of escape at the same distance Relativistic cosmology simply selects a space of positive curvature when $v^2 < v_0^2$, a space of negative curvature when $v^2 > v_0^2$ The space chosen is of course totally different from the space commonly used in physics or in our kinematic solutions All these spaces are simply *maps* of various kinds But whilst the kinematic method uses a map identical with the map used in everyday life, best suited to our other experiences, relativistic cosmology uses a map in which a given particle is given fixed coordinates, and then secures motion for the particle by making the map expand In all cases there is no real entity, 'physical space', and accordingly no physical quantity 'the curvature of space' describing a physical property of any such entity When writers on relativistic cosmology sometimes state that we are unfortunately ignorant of the true curvature of space, all they mean is that we cannot as yet determine whether the nebular velocities exceed, equal, or are less than the parabolic velocity of escape, reckoned locally with respect to a given observer The 'curvature of space', which has been endowed with such mystery, is merely a mathematical device for describing on both the large and the small scale what is equally well described on the small scale (locally) in Newtonian dynamical terms, and what is formally describable on the large scale in the same terms

456. Our analysis thus throws a flood of light on the cumbrous mathematical paraphernalia of 'general' relativity The deduction of equations (45) and (47) by the methods of the tensor calculus from the field equations (50) is a tedious process, involving the calculation of numerous Christoffel symbols In some way which requires deeper

investigation, the simple Newtonian calculus gets to the same destination with incomparably less labour The present position awaits the insight of an investigator who will tell us why the Newtonian analysis reproduces the results of general relativity obtained via the tensor calculus The parallelism must go deep down into the fundamentals of relativity and dynamics For the relativistic cosmological equations (45) and (47) are formally independent of the value of c, for they reduce to (36) and its predecessor Thus the whole formalism of relativistic cosmology is strictly independent of the value assigned to c, i e of the ratio of our usual length and time units The *interpretation* of the formalism depends intimately on the value of c, as we shall see in the next chapter In some way Newtonian dynamics and gravitation embody a formalism completely and absolutely equivalent to the relativistic formalism, the two schemes differ in their interpretation of the formalism in terms of observations As stated earlier, it would be of the first importance to ascertain whether a similar formalism exists in other abstract gravitational situations

457. In the meantime the parallelism of the Newtonian and relativistic formalisms can be used to interpret the relativistic universes employing expanding spaces of positive or negative curvatures, just as we interpreted the expanding space of zero curvature For example, in the Newtonian case in which $A > 0$, each particle is steadily decelerated relative to the chosen observer, and comes ultimately to a constant limiting velocity, given by

$$\lim_{t\to\infty} v = A^{\frac{1}{2}}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}[M(r)]^{\frac{1}{3}}$$

In the corresponding hyperbolic relativistic case ($A > 0$, $k < 0$) the corresponding property is that the observed Doppler displacement for any assigned particle steadily decreases and tends to a constant non-zero limit as the epoch of observation advances The particle thus appears to decelerate, in the experience of the observer The details of the proof of this are worth consideration

458. If a light-signal leaves a particle at local time t' and arrives at the observer at his time t_2, then since $ds = 0$ along the light ray we have by (48)

$$\int_{t}^{t_2} \frac{c\,dt}{R(t)} = \int d\sigma = [\sigma] \tag{52}$$

For a particle of fixed coordinates in the space defined by (48), $[\sigma]$ is constant, and so, as before, the corresponding epochs $t'+dt'$ and t_2+dt_2 for a neighbouring signal are given by

$$\frac{dt_2}{R(t_2)}-\frac{dt'}{R(t')}=0, \tag{53}$$

whence the Doppler shift coefficient s is given by

$$s=\frac{dt_2}{dt'}=\frac{R(t_2)}{R(t')}>1, \tag{54}$$

since $t_2>t'$ and R is an increasing function of its argument Consequently the Doppler shift coefficient s will decrease for any given particle as the epoch of observation advances provided

$$\frac{d}{dt_2}\left(\frac{R(t_2)}{R(t')}\right)<0,$$

i e provided $$\frac{R'(t_2)dt_2}{R(t_2)}-\frac{R'(t')\,dt'}{R(t')}<0,$$

i e by (53) provided $$R'(t_2)<R'(t')$$

But R is the same function of cosmic time t as the Newtonian distance r is of Newtonian time t (apart from a multiplying constant), in the corresponding Newtonian universe with the same A Now in the Newtonian universe every particle is *decelerated* Hence $d^2r/dt^2<0$, and so $d^2R/dt^2<0$ Hence $R'(t)$ is a decreasing function of t Hence $R'(t_2)<R'(t')$ Thus the Doppler effect decreases with advancing epoch of observation Further, since for large t in the Newtonian case we have $v\sim$ const , therefore $r\sim t\times$ const , and so ultimately in the relativistic case $R\sim t\times$ const Hence ultimately, by (53), $\log(t_2/t')\sim$ const , and so ultimately

$$s\sim\frac{t_2}{t'}\sim\text{const}$$

Hence the Doppler shift decreases to a constant limit This example affords a good illustration of the utility of employing results for a Newtonian universe to secure relativistic results for the corresponding relativistic universe

459. Relation (41) can be integrated in finite terms The form depends on the sign of A It is clear that the density ρ, and so the local density near the observer, as a function of t, depends on the nature of A There is thus no unique local density, for given t, in either the general Newtonian solution or the general relativistic solution We

are led in fact to a continuum of solutions defined by the value assigned to A. This is in marked contrast with the kinematic solution, in which the local density is $1/\frac{4}{3}\pi\gamma t^2$ uniquely, independent of the arbitrary constant B defining the solution. We shall examine in the next chapter the cases A small and A large, observed at epochs of small or large t_2.

In all cases it should be noted that the constant of integration b is without physical significance. It never appears in any relation relating observable quantities. The constant A alone possesses a physical significance. For example, the assigned curvature, k/R^2, is simply $-A\rho^{\frac{2}{3}}/c^2$, by (46) and (35).

Modification for the introduction of the cosmical constant 'λ'

460. A solution of the cosmological problem on Newtonian principles is also possible when a term proportional to r is added to the Newtonian accelerations. Thus if we write the Newtonian equation of motion as

$$\frac{Dv}{Dt} = -\gamma\frac{M(r)}{r^2}+\tfrac{1}{3}\lambda c^2 r, \tag{28'}$$

the analysis can be conducted as before and we find

$$v = r[\tfrac{8}{3}\pi\gamma\rho+A\rho^{\frac{2}{3}}+\tfrac{1}{3}\lambda c^2]^{\frac{1}{2}}, \tag{39'}$$

$$t = \int\limits_0^\theta \frac{\theta^{\frac{1}{2}}\,d\theta}{[\frac{8}{3}\pi\gamma+A\theta+\frac{1}{3}\lambda c^2\theta^3]^{\frac{1}{2}}}, \tag{43'}$$

where

$$\rho = 1/\theta^3.$$

Further, the Newtonian differential equations reduce to

$$\left(\frac{1}{Rc}\frac{dR}{dt}\right)^2+\frac{k}{R^2} = \tfrac{1}{3}\kappa\rho+\tfrac{1}{3}\lambda, \tag{45'}$$

$$\frac{2}{R}\frac{d^2R}{c^2\,dt^2}+\left(\frac{1}{R}\frac{dR}{c\,dt}\right)^2+\frac{k}{R^2} = \lambda, \tag{47'}$$

where

$$k = -\frac{Ab^{\frac{2}{3}}}{c^2}, \qquad \rho = \frac{b}{R^3}.$$

These are identical in form with the equations of relativistic cosmology for $p = 0, \lambda \neq 0$.† It follows as before that there is a complete correspondence between the relativistic universes and the Newtonian universes as modified by the λ-term, and that the local phenomena in the two systems, as observable, are identical.

† Robertson or de Sitter, *loc. cit.* These equations are due essentially to Friedmann, *Zeits. für Phys.*, **10**, 377, 1922; **21**, 326, 1924, and to Lemaître, *Ann. Soc. Sci. Bruxelles*, **47** A, 49, 1927; *M.N., R.A.S.*, **91**, 483, 1931.

461. It may be remarked that an additive acceleration proportional to r is the only permissible modification of the Newtonian inverse square law which is compatible with the satisfaction of the cosmological principle, any other modification would make it impossible for every observer to calculate accelerations, etc, as though he were central This fact sheds an altogether deeper significance on the cosmological principle than its mere adoption as a selective definition of the systems to be considered It suggests a reason why the building-up of relativistic kinematic systems satisfying the cosmological principle reproduces phenomena locally reconcilable with an assumed inverse square law and a cosmical repulsion, though in the kinematic treatment both appear as mere arbitrary resolutions, by the observers, of observable accelerations into the effects of action at a distance, and so are artificial

462. There is no theoretical or observational justification for the adoption of a cosmic term λr as expressing an objective law of nature General relativity was driven to introduce it in an endeavour to secure itself from the difficulties imposed by its own limitations In a kinematical treatment these limitations disappear, and there is no need either to assume, or not to assume, the existence of a 'cosmical constant' λ, any more than to assume the existence or non-existence of a 'constant of gravitation' The accelerations undergone by particles in the presence of other particles have been shown to be capable of discussion, in systems satisfying the cosmological principle, without recourse to an assumed existence of 'laws of nature' Further, the field equations of 'general' relativity, even if objective laws of nature are assumed, are equally compatible with any value of λ, and general relativity from its own principles gives no reason for supposing λ to be small It appeals, in fact, empirically to the observed smallness or non-existence of any λ-term in local gravitational situations, thus revealing an essentially empirical constituent in the framework of 'general' relativity It has become fashionable of late to consider it as more satisfactory on general or aesthetic grounds to take $\lambda = 0$, but this procedure is equally empirical There is no more fundamental justification for assigning λ the value zero than for assigning it any other value In having to face the necessity for a decision as to the presence or absence of λ, 'general' relativity encounters its limitations We have shown that at this point progress comes not by trying

to generalize 'general' relativity but by going to the underlying kinematic principles on which any relativity must be based, and re-formulating the theory

In the further comparisons, given in the next chapter, between kinematic and 'general' relativity, we shall only treat in detail the case $\lambda = 0$, not on aesthetic or scientific grounds but by way of illustration The main features of the case $\lambda \neq 0$ can then be inferred by general arguments

XVII

COMPARISON BETWEEN THE SIMPLE KINEMATIC SYSTEM AND THE SYSTEMS OF RELATIVISTIC COSMOLOGY

463. We begin by examining the world-model of Einstein and de Sitter, namely that described in terms of flat expanding space, and constructing its world-picture

The universe of Einstein and de Sitter

464 In this system the metric is given by

$$ds^2 = c^2\,dt^2 - [R(t)]^2(dx^2+dy^2+dz^2), \tag{1}$$

and the solution of the associated field equations has been shown to be given by

$$R(t) = bt^{\frac{2}{3}}, \qquad v = \frac{2}{3}\frac{r}{t}, \qquad \rho = \frac{1}{6\pi\gamma t^2} \tag{2}$$

In these formulae b is arbitrary, and γ is a constant We proceed to find, for reasons explained in Chapter XV, (i) the behaviour of the Doppler shift s for any particle of fixed coordinates as a function of the epoch of observation t_2, (ii) the number of particles, inside a solid angle $d\omega$, possessing at epoch of observation t_2 Doppler shifts lying between s and $s+ds$, (iii) the local time t' on the particle observed as a function of s and the epoch of observation t_2

465. Transforming the space-section of (1) to polar coordinates r, ψ, ϕ, it becomes

$$ds^2 = c^2\,dt^2 - [R(t)]^2(dr^2+r^2\,d\psi^2 + r^2\sin^2\psi\,d\phi^2),$$

where $\quad x = r\cos\psi, \qquad y = r\sin\psi\cos\phi, \qquad z = r\sin\psi\sin\phi$

The element of volume in the three-space near r, ψ, ϕ is accordingly

$$[R(t)]^3 r^2\,dr\sin\psi\,d\psi d\phi$$

or

$$[R(t)]^3 r^2\,dr d\omega,$$

where $d\omega$ is the element of solid angle The mass enclosed in the element, measured locally, is

$$\rho[R(t)]^3 r^2\,dr d\omega,$$

and if m_0 is the rest-mass of a particle, the number of particles in this element is

$$\frac{\rho}{m_0}[R(t)]^3 r^2\,dr d\omega$$

Substituting from (2) for ρ and $R(t)$, we see that t cancels, and the number comes out as

$$\frac{b^3}{6\pi\gamma m_0} r^2\,dr d\omega \tag{3}$$

For an element $d\omega$ of the shell of fixed coordinates r, $r+dr$, this number is a constant, as it should be, independent of the epoch

466 We now calculate the Doppler shift coefficient s for these particles If a light-signal leaves the particle defined by r, ψ, ϕ at local time t' and arrives at the observer at this observer's time t_2, then since $ds = 0$ along the path,

$$r = \int dr = \int_{t'}^{t_2} \frac{c\,dt}{R(t)} = \frac{3c}{b}(t_2^{\frac{1}{3}} - t'^{\frac{1}{3}}) \tag{4}$$

For two neighbouring light-signals, from the same source to the same observer, we have

$$0 = \frac{dt_2}{R(t_2)} - \frac{dt'}{R(t')},$$

or

$$s = \frac{dt_2}{dt'} = \frac{R(t_2)}{R(t')} = \left(\frac{t_2}{t'}\right)^{\frac{2}{3}} \tag{5}$$

Hence by (4)

$$r = \frac{3c}{b} t_2^{\frac{1}{3}}(1 - s^{-\frac{1}{2}}), \tag{6}$$

or

$$s = \frac{1}{\left(1 - \dfrac{br}{3ct_2^{\frac{1}{3}}}\right)^2} \tag{7}$$

This gives the behaviour of s as a function of t_2 for a given particle of fixed coordinate r, and is the answer to (i)

467. Differentiating (6) at a fixed epoch of observation t_2, we see that a range dr in the coordinate r corresponds to a range ds in the Doppler shift, where

$$dr = \frac{3ct_2^{\frac{1}{3}}}{2b} s^{-\frac{3}{2}}\, ds \tag{8}$$

Inserting for r and dr from (6) and (8) in (3), we see that the number of particles inside a solid angle $d\omega$, with Doppler shifts between s and $s+ds$, is

$$\frac{9}{4\pi m_0} \frac{c^3 t_2}{\gamma} \frac{(1-s^{-\frac{1}{2}})^2}{s^{\frac{3}{2}}} ds d\omega \tag{9}$$

This is the answer to (ii)

468. The answer to (iii) is given immediately by (5) in the form $t' = t_2/s^{\frac{3}{2}}$ We now tabulate our three world-picture relations (relations between observable quantities) in the forms

$$s = \frac{1}{\left(1 - \dfrac{br}{3ct_2^{\frac{1}{3}}}\right)^2}, \tag{i$'$}$$

$$N\,ds\,d\omega = \frac{9}{4\pi m_0}\frac{c^3 t_2}{\gamma}\frac{(1-s^{-\frac{1}{2}})^2}{s^{\frac{3}{2}}}\,ds\,d\omega, \tag{ii$'$}$$

$$t' = t_2/s^{\frac{3}{2}} \tag{iii$'$}$$

469. These may now be compared directly with the three corresponding formulae for the simple kinematic system obtained at the end of Chapter XV and called there (i), (ii), (iii) We see at once that all three relationships are essentially different for the Einstein-de Sitter system as compared with the simple kinematic system We take them in turn

(i) In the simple kinematic system, s for any given particle is constant In the Einstein-de Sitter system (parabolic case), by (i′) s decreases as t_2, the epoch of observation, advances, as $t_2 \to \infty$, $s \to 1$, and the particle tends ultimately to possess a zero Doppler shift Thus in the parabolic system, every particle appears ultimately to come to rest

(ii) In the simple kinematic system, N is a constant multiple of $(s^2-1)^2/s^3$, this tends to zero as $s \to 1$, and $\to \infty$ as $s \to \infty$ Thus the count of nebulae must show a constant population when classified in terms of Doppler shift s In the Einstein-de Sitter system, N is proportional first to t_2, secondly to $(1-s^{-\frac{1}{2}})^2/s^{\frac{3}{2}}$ The observable population similarly classified therefore increases with advancing epoch of observation t_2, for $s = 1$ it is zero, and as $s \to \infty$ it tends to zero

(iii) In the simple kinematic system, t' is proportional to t_2 and inversely proportional to s, in the Einstein-de Sitter system t' is also proportional to t_2, but is inversely proportional to $s^{\frac{3}{2}}$

This is sufficient to show that the kinematic system (described in flat space) appears to observation fundamentally different from the Einstein-de Sitter system (described in flat expanding space) The two structures are thus entirely different

470 The Einstein-de Sitter system possesses certain other properties of a striking character which compel us in the end to reject it Consider formula (ii′) The total number of particles observable at epoch t_2, of all possible Doppler shifts, is

$$\int d\omega \int_{s=1}^{\infty} N\,ds = \frac{9c^3 t_2}{\gamma m_0}\int_1^{\infty}\frac{(1-s^{-\frac{1}{2}})^2}{s^{\frac{3}{2}}}\,ds = \frac{6c^3 t_2}{m_0\gamma} \tag{10}†$$

† Formula (10) bears a superficial resemblance to the 'extrapolated homogeneous population' of the simple kinematic system, $\frac{4}{3}\pi B$, for this quantity has been shown in

Thus $6c^3t_2/\gamma$ is a measure of the total mass of the system observable at epoch t_2 This is of the order of all estimates of the 'mass of the universe' But here the mass observable increases as t_2 increases, being proportional to t_2, for by definition γ is a constant In fact, as $t_2 \to \infty$ the total mass observable $\to \infty$ Where does the new mass enter the field of observation?

471 So far we have tacitly ignored the possibility of negative values for t' If we permit them, (5) permits $t' = -t_2 s^{-\frac{2}{3}}$ and (6) gives $r = (3ct_2^{\frac{1}{3}}/b)(1+s^{-\frac{1}{3}})$ Positive values of t' then correspond to $1 \leqslant s < \infty$, $0 < r < 3ct_2^{\frac{1}{3}}/b$, negative values of t' to $\infty > s > 0$, $r > 3ct_2^{\frac{1}{3}}/b$ At $r = 3ct_2^{\frac{1}{3}}/b$ we have $s = \infty$ and $t' = 0$

The resulting world-picture at t_2 would be as follows † The observer sees, surrounding himself, a spherical domain of receding particles, extending from $s = 1$ to $s = \infty$ At $s = \infty$, the particles are viewed at their natural time-zero, $t' = 0$, these particles are receding with the velocity of light, and therefore just invisible Outside this domain would appear a zone of receding particles, extending from $s = \infty$ to $s = 1$, observed at negative values of t' At $s = 1$ the particles would appear at rest relative to the observer Outside these again, and extending indefinitely, would appear a domain of *approaching* particles, $1 > s > 0$, also viewed at negative local times At indefinitely large distances, $s \to 0$ and the approach velocity tends to that of light

The trajectory of any individual particle, of fixed coordinate r, would be as follows At great distances from the observer ($t' \sim -\infty$), the particle would be approaching with the speed of light, but subject to a repulsive acceleration The approach-velocity first reduces to zero ($s = 1$), then changes sign and becomes a recession velocity which increases to that of light ($s = \infty$), acquired at epoch of observation $t_2 = (br/3c)^3$, lastly decreases to zero ($s \to 1$) as t_2, $t' \to +\infty$ At the epoch $t_2 = (br/3c)^3$, t' is zero and the acceleration changes sign from repulsion to attraction

Chapter XIV to be equal to $c^3t_2/m_0\gamma$ In both cases, c^3t/γ is a fundamental mass associated with the description of the system But in (10) γ is a constant, whilst in the corresponding formula for the simple kinematic system $\gamma \propto t_2$, γ being here merely a convenient measure of the acceleration of an inter galactic free particle, attributed to the matter between the observer and the free particle but actually arising from the presence of the infinity of particles present

† Cf McCrea, *Zeits für Astrophys*, **9**, 290, 1935 I am indebted to Dr McCrea for discussion on the subject of negative t'

472. Since here $\lambda = 0$, this change of sign appears inexplicable on physical grounds But there are two other reasons for rejecting this interpretation of the model First, whatever luminosity is attributed to an individual nebula for $t' \sim -\infty$, the density-distribution at great distances (where $s \sim 0$) is such that the total sky-brightness would be infinite The model violates in fact Seeliger's well-known condition for finite sky-brightness †

Secondly, as shown in Chapter IV, it is highly improbable that observation could ever disclose a contracting system Yet to permit negative values for t' is to suppose that a contracting phase of the system is open to observation at t_2 Further, since the local density is infinite at $t' = 0$, it is absurd to suppose that we can 'see beyond' this fundamental singularity If negative values of t' are allowed, the world-picture is in fact entirely fantastic

For these reasons we disallow negative values of t' and reject the associated interpretation of the Einstein-de Sitter model We can only rescue the model by supposing that, in it, particles are actually created on the expanding frontier $s = \infty$, $t' = 0$, each newly created particle r is then first open to observation at the epoch $t_2 = (br/3c)^3$ The model is now free from internal contradiction, but only at the cost of involving the creation of matter within the actual temporal experience of the observer at the origin

473. Suppose the observer at the origin builds a very big telescope Let him expose it on the sky, at epoch t_2 Since by (10) the total number of particles observable at t_2 is finite, he must see a speckled background containing only a finite number of luminous objects But as he watches he will see fresh particles appear, at first with very small luminosities and very large Doppler shifts, at the maximum distance of any observable particle, as he watches further, each new particle will brighten and decrease in Doppler shift, ultimately fading again and acquiring in the limit a Doppler shift zero ($s = 1$) He actually *sees* the process of creation going on Each particle is born at its own time-zero, and the observer at the origin would see the whole early history of the particle, from $t' = 0$ on

474. This is what the mathematics predicts The mathematics is responsible for this creation of matter, and it does so because it is

† The 'star ratio' tends to a constant See Russell, Dugan, and Stewart, *Astronomy*, **2**, 810

set an otherwise impossible task. It is asked to produce a finite number of observable objects, given by (10), such that each one is central in the field of the remainder. It only achieves this object by *creating* fresh particles beyond each given particle as fast as they are required, and it brings them to birth with the velocity of light. The frontier of limiting range of observability moves onward with the speed of light, it contains always the particles just being created, and it leaves in its wake a spray of decelerating newly created particles.

Since creation of matter within the epoch covered by our own experiences is contrary to our experience, we are bound now to reject the Einstein-de Sitter model. It carries out the task assigned to it by relativity and the requirements of the cosmological principle, but only at the cost of violating experience. Matter is conserved within the expanding light-sphere, but created at its frontier.

475. The case is entirely different for the kinematic model. Here the field of vision contains at every epoch of observation an infinity of particles already created, merging into invisibility at great distances approaching ct, but forming an irresolvable continuous background. There is no 'most distant particle'. The very distant particles possess arbitrarily early local times t' in the present experience (t_2) of the observer at the origin (any arbitrary particle), but t', however small, is always positive. Creation is always a thing of the past, even when only just a thing of the past. Creation is here always pre-experiential, a transcendental *prior* singularity. There is 'all the difference in the world' between an act of creation that no observer can experience or witness, and acts of creation occurring within the temporal experience of each observer. In the Einstein-de Sitter model, every particle is definitely created at some directly calculable epoch in the experience of any given observer, the system creates itself as it expands, instead of displaying, as does the kinematic system, the expansion of an already created system.

476. Investigators have long looked for a criterion which would enable them to choose between the different models proposed by general relativity. We have at last found such a criterion. We reject systems which imply the creation of matter within the finite experience of an observer. We accordingly reject the Einstein-de Sitter model and retain the kinematic model. The Einstein-de Sitter model is not a mass-conserving model, it implies the creation of matter at its

expanding frontier of visibility The kinematic model is, on the other hand, strictly mass-conserving

We now turn to hyperbolic universes, or expanding relativistic universes of negative curvature

Expanding spaces of negative curvature

477. These systems have been seen in Chapter XVI to be defined by the metric

$$ds^2 = c^2\,dt^2 - \frac{[R(t)]^2(dx^2+dy^2+dz^2)}{[1+\frac{1}{4}k(x^2+y^2+z^2)]^2}, \tag{11}$$

in which R is a function of t defined by the relations

$$\rho = \frac{b}{R^3} = \frac{1}{\theta^3}, \qquad k = -\frac{Ab^{\frac{2}{3}}}{c^2}, \tag{12}$$

$$t = \int\limits_0^\theta \frac{\theta^{\frac{1}{2}}\,d\theta}{(\frac{8}{3}\pi\gamma+A\theta)^{\frac{1}{2}}}. \tag{13}$$

We shall examine in detail the case $A > 0$

478. In this case k is negative Transform the spatial section of the metric by the well-known substitution

$$x = \frac{2}{(-k)^{\frac{1}{2}}}\tanh\tfrac{1}{2}\chi\cos\psi, \qquad y = \frac{2}{(-k)^{\frac{1}{2}}}\tanh\tfrac{1}{2}\chi\sin\psi\cos\phi,$$

$$z = \frac{2}{(-k)^{\frac{1}{2}}}\tanh\tfrac{1}{2}\chi\sin\psi\sin\phi \tag{14}$$

Then the metric (1) becomes

$$ds^2 = c^2\,dt^2 - \frac{[R(t)]^2}{(-k)}[d\chi^2+\sinh^2\chi(d\psi^2+\sin^2\psi\,d\phi^2)] \tag{15}$$

In the three-space of this metric the volume element is

$$\frac{R^3}{(-k)^{\frac{3}{2}}}\sinh^2\chi\sin\psi\,d\chi d\psi d\phi,$$

or

$$\frac{R^3}{(-k)^{\frac{3}{2}}}\sinh^2\chi\,d\chi d\omega, \tag{16}$$

where $d\omega$ is the spatial solid angle The proper mass enclosed between the spheres χ and $\chi+d\chi$, within the solid angle $d\omega$, is accordingly

$$\frac{R^3\rho}{(-k)^{\frac{3}{2}}}\sinh^2\chi\,d\chi d\omega$$

If m_0 is the proper mass of a particle or nebula, the number of nebulae in the element of shell $d\chi d\omega$ is thus

$$\frac{R^3\rho}{m_0(-k)^{\frac{3}{2}}}\sinh^2\chi\, d\chi d\omega$$

Substituting for R, ρ, and k from (12), this number is found to be

$$\frac{c^3}{m_0 A^{\frac{3}{2}}}\sinh^2\chi\, d\chi d\omega \tag{17}$$

This, as it should be, is a constant independent of the time, and independent also of the purely conventional number b

479 We now consider Doppler effects Let a light-signal leave the particle (χ, ψ, ϕ) at local time t', and arrive at the observer at the origin of χ at his local time t_2, t_2 is his epoch of observation Along this light-track, $ds = 0$ and ψ and ϕ are constant Hence by (15) $c\,dt = -R(t)d\chi/(-k)^{\frac{1}{2}}$ and so the coordinate χ defining the distant particle considered is given by

$$\chi = -\int\limits_{\chi}^{0} d\chi = (-k)^{\frac{1}{2}}\int\limits_{t}^{t} \frac{c\,dt}{R(t)} \tag{18}$$

Substituting from (12) for k and R, we have

$$\chi = A^{\frac{1}{2}}\int\limits_{t'}^{t_2} \frac{dt}{\theta(t)},$$

independent again of b as it should be By (13) this is

$$\chi = A^{\frac{1}{2}}\int\limits_{\theta}^{\theta_2} \frac{d\theta}{\theta^{\frac{1}{2}}(\frac{8}{3}\pi\gamma+A\theta)^{\frac{1}{2}}}. \tag{19}$$

Here θ' and θ_2 are respectively $\rho'^{-\frac{1}{3}}$, $\rho_2^{-\frac{1}{3}}$, where ρ', ρ_2 are the densities at times t' and t_2

By (18), for two neighbouring light-signals from the particle of given coordinate χ,

$$\frac{dt_2}{R(t_2)} - \frac{dt'}{R(t')} = 0,$$

whence the Doppler shift coefficient s for the particle at (χ, ψ, ϕ) as observed by the observer at $\chi = 0$ at his epoch t_2 is given by

$$s = \frac{dt_2}{dt'} = \frac{R(t_2)}{R(t')} = \frac{\theta_2}{\theta'} \tag{20}$$

But by (13) the epoch t_2 is given in terms of θ_2 by

$$t_2 = \int\limits_0^{\theta} \frac{\theta^{\frac{1}{2}}\, d\theta}{(\frac{8}{3}\pi\gamma + A\theta)^{\frac{1}{2}}}, \tag{21}$$

whilst the epoch t' is given in terms of θ_2 by

$$t' = \int\limits_0^{\theta_2/s} \frac{\theta^{\frac{1}{2}}\, d\theta}{(\frac{8}{3}\pi\gamma + A\theta)^{\frac{1}{2}}}, \tag{22}$$

on putting $\theta' = \theta_2/s$ by (20). Equation (19) may be rewritten as

$$\chi = A^{\frac{1}{2}} \int\limits_{\theta/s}^{\theta} \frac{d\theta}{\theta^{\frac{1}{2}}[\frac{8}{3}\pi\gamma + A\theta]^{\frac{1}{2}}} \tag{23}$$

480. The constant A fixes the model under examination and is supposed known, if the actual universe followed this model, A could be found observationally by comparing the recession-ratio v/r with the local mean density ρ, by formula (39), Chapter XVI. A being known, or given, equation (21) then fixes θ_2 as a function of the epoch of observation t_2.† Equation (23) then determines the value of s at epoch t_2 for a particle of fixed coordinate χ. Equation (22) determines the local time t' on the nebula defined by s as observed at epoch of observation t_2.

481. We must now determine the number of particles or nebulae with Doppler shifts lying in a given range $(s, s+ds)$ as observed at epoch t_2. If this range corresponds to the range $(\chi, \chi+d\chi)$ of the coordinate χ, then differentiating (23), keeping t_2 and so θ_2 fixed, we have

$$d\chi = \frac{A^{\frac{1}{2}}\theta_2^{\frac{1}{2}}}{(\frac{8}{3}\pi\gamma + A\theta_2/s)^{\frac{1}{2}}} \frac{ds}{s^{\frac{3}{2}}} \tag{24}$$

Substituting for χ and $d\chi$ from (23) and (24) in (17), the number $N\, ds d\omega$ of particles at epoch of observation t_2 with Doppler shifts lying between s and $s+ds$, inside the solid angle $d\omega$, is

$$N\, ds d\omega = \frac{c^3}{m_0 A}\left[\sinh\left\{A^{\frac{1}{2}} \int\limits_{\theta_2/s}^{\theta_2} \frac{d\theta}{\theta^{\frac{1}{2}}(\frac{8}{3}\pi\gamma + A\theta)^{\frac{1}{2}}}\right\}\right]^2 \frac{\theta_2^{\frac{1}{2}}}{(\frac{8}{3}\pi\gamma + A\theta_2/s)^{\frac{1}{2}}} \frac{ds}{s^{\frac{3}{2}}} d\omega \tag{25}$$

482. Essentially, equations (23), (25), (22) provide the observable relations we have called (i), (ii), and (iii) respectively, for the model

† Conversely, knowledge of the local density ρ_2 fixes θ_2, and then (21) fixes t_2.

defined by A, with θ_2 given in terms of t_2 by (21) To deal with these formulae we now make a transformation which leads to the isolation of a new dimensionless parameter determining the state of affairs

483. We have previously seen, in discussing the parabolic model ($A = 0$), that the density ρ_2 is given in that case by $1/6\pi\gamma t_2^2$ This suggests using the quantity $1/6\pi\gamma t_2^2$ again in the present case ($A \neq 0$) as a standard of density We accordingly introduce a variable of integration x and a particular value of it x_2, defined by the relations

$$\theta = (6\pi\gamma t_2^2)^{\frac{1}{3}}x, \qquad \theta_2 = (6\pi\gamma t_2^2)^{\frac{1}{3}}x_2 \tag{26}$$

The variable x is dimensionless Then (21) becomes

$$1 = \frac{3}{2}\int\limits_0^{x_2} \frac{x^{\frac{1}{2}}\,dx}{(1+\alpha x)^{\frac{1}{2}}}, \tag{27}$$

where

$$\alpha = \frac{9}{4}\frac{At_2^{\frac{4}{3}}}{(6\pi\gamma)^{\frac{1}{3}}}, \tag{28}$$

α is a new dimensionless parameter Equation (27) defines x_2 as a function of α, and so of t_2 Similarly (22) becomes

$$t' = \frac{3}{2}t_2\int\limits_0^{x_2/s} \frac{x^{\frac{1}{2}}\,dx}{(1+\alpha x)^{\frac{1}{2}}}, \tag{29}$$

(23) becomes

$$\chi = \alpha^{\frac{1}{2}}\int\limits_{x_2/s}^{x_2} \frac{dx}{x^{\frac{1}{2}}(1+\alpha x)^{\frac{1}{2}}}, \tag{30}$$

and (25) becomes

$$N\,dsd\omega = \frac{9c^3t_2}{16\pi m_0\gamma}\left[\frac{1}{\alpha^{\frac{1}{2}}}\sinh\int\limits_{x_2/s}^{x_2} \frac{\alpha^{\frac{1}{2}}\,dx}{x^{\frac{1}{2}}(1+\alpha x)^{\frac{1}{2}}}\right]^2 \frac{x_2^{\frac{1}{2}}}{(1+\alpha x_2/s)^{\frac{1}{2}}}\frac{ds}{s^{\frac{3}{2}}}d\omega \tag{31}$$

The local density ρ_2 at epoch of observation t_2 is given by

$$\rho_2 = \frac{1}{\theta_2^3} = \frac{1}{6\pi\gamma t_2^2 x_2^3} \tag{32}$$

484. We notice, first, the emergence of the characteristic factor c^3t_2/γ in (31), common to all cosmologies We notice, secondly, that all relations appear in terms of the single dimensionless parameter α

485. It is clear without further discussion that for any given A, at any given epoch t_2, the observable relations now found differ from

the corresponding relations for the simple kinematic system But two special limiting cases of these formulae are of great interest

486. *Case* (*a*) $\alpha \sim 0$ When α is *small*, (27) becomes approximately

$$1 \sim x_2^{\frac{3}{2}},$$

(29) becomes
$$t' \sim t_2\left(\frac{x_2}{s}\right)^{\frac{3}{2}} \sim \frac{t_2}{s^{\frac{3}{2}}}, \qquad \text{(iii}''\text{)}$$

(30) becomes
$$\chi \sim 2\alpha^{\frac{1}{2}}x_2^{\frac{1}{2}}(1-s^{-\frac{1}{2}}) \sim 2\alpha^{\frac{1}{2}}(1-s^{-\frac{1}{2}}), \qquad \text{(i}''\text{)}$$

(31) becomes
$$N\,dsd\omega \sim \frac{9}{4\pi m_0}\,\frac{c^3 t_2}{\gamma}\,\frac{(1-s^{-\frac{1}{2}})^2}{s^{\frac{3}{2}}}\,dsd\omega, \qquad \text{(ii}''\text{)}$$

(32) becomes
$$\rho_2 \sim 1/6\pi\gamma t_2^2 \qquad (33)$$

Comparison of these with formulae (i′), (ii′), (iii′) given (§ 468) in this chapter shows that for α small, the model for $A > 0$ coincides in observable properties with the parabolic Einstein-de Sitter model $A = 0$ (The coordinate χ is equal to $(br/c)A^{\frac{1}{2}}(6\pi\gamma)^{-\frac{1}{3}}$) Now α is small when the product $At_2^{\frac{2}{3}}$ is sufficiently small This occurs when A is sufficiently small, for any t_2, and when t_2 is sufficiently small, for any A It is not surprising that the case $A \sim 0$ for any t_2 is approximately represented by the case $A = 0$ But it is rather unexpected that for *any* A, for t_2 sufficiently small, the hyperbolic universe is approximately coincident with the parabolic universe We may say that every hyperbolic universe, or universe of negative curvature, starts life and behaves in its youth as a parabolic universe, or universe of zero curvature This result is not really relevant to the main purpose of this book, but I have inserted it in order to illustrate the present method and also to demonstrate the highly conventional character of the 'curvature' of space in relativistic phraseology The two models $A = 0$ and $A > 0$ have totally different space-curvatures for their relativistic descriptions, yet give to observation the same appearance for t_2 small Identical observational appearances are seen here to be compatible with totally different geometries

487 *Case* (*b*) α *large* When α is large, (27) becomes

$$x_2 \sim \tfrac{2}{3}\alpha^{\frac{1}{2}}, \qquad (34)$$

(29) becomes
$$t' \sim \frac{3}{2}\,\frac{t_2}{\alpha^{\frac{1}{2}}}\left(\frac{x_2}{s}\right) \sim \frac{t_2}{s}, \qquad \text{(iii}'''\text{)}$$

(30) becomes
$$\chi \sim \int_{x/s}^{x_2} \frac{dx}{x} \sim \log s, \qquad \text{(i}'''\text{)}$$

whence $$\sinh\chi \sim \sinh(\log s) = \frac{s^2-1}{2s}$$

Employing the latter relation in (31), we have

$$N\,dsd\omega \sim \frac{9c^3t_2}{16\pi m_0\gamma}\,\frac{1}{\alpha},\,\frac{(s^2-1)^2}{4s^3}\,dsd\omega,$$

or, using the value of α given by (28),

$$N\,dsd\omega \sim \frac{c^3}{4m_0A^{\frac{3}{2}}}\,\frac{(s^2-1)^2}{s^3}\,dsd\omega \tag{ii$'''$}$$

Lastly, by (32) and (34), $$\rho_2 \sim \frac{1}{A^{\frac{3}{2}}t_2^3} \tag{35}$$

488. Comparison of (i$'''$), (ii$'''$), (iii$'''$) with (i), (ii), (iii) of the end of Chapter XV shows that when α is large, the observable properties of the hyperbolic relativistic models ($A > 0$) become approximately the same as those of the observable properties of the simple kinematic model In particular, we notice from (i$'''$) that for any fixed particle (fixed χ) the Doppler shift s tends to a constant, and so the particle has a constant velocity in the coordinates and space of ordinary physics, (iii$'''$) shows that the local time on the distant particle, as observed at epoch of observation t_2, is the same function of t_2 and s as in the kinematic case, and (ii$'''$) shows that counting of particles of given range of Doppler shift gives the same result

489 Now by (28), α is large either when t_2 is large, for given A, or when A is large, for given t_2 Thus, for any 'curvature' defined by A, the hyperbolic relativistic model tends to pass into the kinematic model as the epoch of observation advances, and the two systems are approximately alike to observation, at any epoch of observation, if A is chosen sufficiently large We notice from (35) and from the formula for the local density at t_2 in the kinematic case, namely $\rho_2 = m_0 B/c^3t_2^3$, that the local density (near the observer) follows the same law of the inverse cube of the time

490 The former of these results, that the two systems pass into coincidence as $t_2 \to \infty$ and accordingly $\rho_2 \to 0$, was obtained independently by H P Robertson† and by W O Kermack and W H McCrea ‡ These authors used a method quite different from that here employed, and they did not show that the approximate equivalence depends not primarily on the size of t_2 but on the size of the product $At_2^{\frac{2}{3}}$, they did not isolate the parameter α Consequently

† *Zeits fur Astrophys*, **7**, 153, 1933 ‡ *M N, R A S*, **93**, 519, 1933

they did not notice the other equally interesting result that the approximate equivalence holds also for any t_2, for A large enough Their method depended essentially on showing that the metric (11) is equivalent to the metric $ds^2 = c^2\,dt^2-(dx^2+dy^2+dz^2)$ for t large, this follows from the fact that the density tends to zero The point is that a hyperbolic relativistic model is not completely defined by the *form* of its metric, but depends on a constant A as well

491 The approximate equivalence of the hyperbolic relativistic model defined by A (>0) to the kinematic model defined by B holds when α is large and when, in addition,

$$\frac{1}{A^{\frac{3}{2}}} = \frac{B}{m_0\,c^3} \tag{36}$$

This follows by comparing formulae (11) and (11‴) or by comparing the local densities ρ_2 Combining the conditions

$$\frac{9}{4}\frac{At_2^{\frac{2}{3}}}{(6\pi\gamma)^{\frac{2}{3}}} \gg 1, \qquad \rho_2 = \frac{1}{A^{\frac{3}{2}}t_2^3}, \tag{37}$$

by elimination of t_2, the condition may be thrown into the form

$$\left(\frac{3}{2}\right)^3 \frac{A}{6\pi\gamma} \gg \rho_2^{\frac{1}{3}} \tag{38}$$

This shows that given ρ_2, the local density, we can always choose A so large that the corresponding hyperbolic relativistic universe is approximately the same to observation as the simple kinematic system of the same density near the observer, whatever this density The corresponding value assigned to t_2 is $1/\rho_2^{\frac{1}{3}}A^{\frac{1}{2}}$ and so may be small This approximate identity for *any* value of ρ_2, on appropriate choice of A, was not stated by Robertson or Kermack and McCrea Their result, which has been widely noticed, has led to the belief that the approximate equivalence of the relativistic and kinematic systems holds only for *small* densities, i e for an almost empty universe This belief is false, approximate equivalence can be arranged for any density

492. It is of interest to inquire to what extent the actual universe satisfies the mathematical conditions of equivalence of the two systems If (38) holds, then $A\rho_2^{\frac{2}{3}} \gg \frac{16}{9}\pi\gamma\rho_2$, and hence equation (39), Chapter XVI, requires

$$\frac{v}{r} \gg (\tfrac{8}{3}\pi\gamma\rho_2)^{\frac{1}{2}},$$

or

$$\rho_2 \ll \frac{3}{8\pi\gamma}\left(\frac{v}{r}\right)^2$$

For v/r corresponding to 500 km sec $^{-1}$ per 10^6 parsecs, this gives

$$\rho_2 \leqslant 0\,5 \times 10^{-27} \text{ gramme cm}^{-3}$$

Present estimates of the local mean density assign as a *minimum* value 10^{-30} gramme cm $^{-3}$ Were the actual mean density equal to the present estimate of the minimum density, it would follow from this that *if* the universe could be represented by a hyperbolic relativistic model, then the history of the universe is so advanced in time that it is locally approximately indistinguishable from a simple kinematic system

Rejection of hyperbolic models

493. There are, however, fundamental objections to identifying the universe with a hyperbolic relativistic model, of the same nature as those we encountered in the case of the parabolic Einstein-de Sitter model Let us consider a particle or nebula at the theoretically maximum distance of observability on this model (*any* A), at any epoch of observation t_2 For such a particle or nebula $s = \infty$, and so by (29) $t' = 0$, the local time, as observed, is the natural time-zero † By (30) the coordinate χ of this nebula takes the value

$$\chi_{\max}(t_2) = \int\limits_0^{x_2} \frac{dx}{x^{\frac{1}{2}}\{x+(1/\alpha)\}^{\frac{1}{2}}}, \tag{39}$$

where x_2 is given as a function of α by (27) From (27), as t_2 increases, and α accordingly increases, x_2 increases, and so by (39) $\chi_{\max}(t_2)$ increases Hence as the epoch of observation advances, new nebulae, defined by increasing values of $\chi_{\max}$, come into the theoretically possible range of observation This is equivalent to stating that there is continual creation of matter at the frontier of observation, at the maximum possible distance For every nebula appears at the time-zero in its local history, and so has no pre-history † In the field of a sufficiently powerful telescope, new nebulae would be continually appearing as new spots of light not previously in observable existence Each would appear at some definite calculable epoch, and with a red-shift, at first infinite, then decreasing as the epoch of observation advanced This creation of matter within the temporal experience of the observer is in contradiction with physics, and compels the rejection of the hyperbolic models with $\lambda = 0$

† To allow negative values for t' would be to encounter the infinite-sky-brightness difficulty, as before

494. The fact is that the approximate equivalence of the hyperbolic and kinematic models, under certain conditions, only extends to a definite distance. At 'greater distances', it breaks down. A look-out man inspecting a hyperbolic universe with a big telescope sees a speckled background—a finite number of luminous objects at any one epoch given by integrating (31) from $s = 1$ to $s = \infty$ (the integral is readily seen to converge, and can be evaluated). These luminous points would be visible as discretely separated on a black background. But a look-out man inspecting similarly the kinematic model would see a continuous background, into which merge the luminous objects, infinite in number, with increasing faintness. In the kinematic model no creation of matter occurs within the temporal experience of any observer; the whole system is already created and in existence by the time any observer can observe it. At any given epoch t_2, there is thus a fundamental observational difference between the hyperbolic and kinematic universes. As t_2 increases, a more and more nearly continuous background comes into being in the relativistic world-picture, but it is always essentially resolvable. It converges to the kinematic world-picture as $t_2 \to \infty$, by the creation of matter, but always differs essentially from it.

Thus the kinematic model avoids the difficulties encountered when the significance of the parabolic and hyperbolic systems is considered in detail. It achieves this because it originates in an analysis of possible *experiences*, whilst the relativistic systems originate in a conceptual scheme of relations.

Oscillating universes

495. It would be tedious to consider in similar detail the elliptic relativistic universes ($\lambda = 0$, $A < 0$), usually described as oscillating universes. They, too, begin as observationally indistinguishable from a parabolic, Einstein-de Sitter universe. Just like the hyperbolic universes, they create themselves at first as they expand, by the creation of matter at their expanding frontiers. But the velocity of any particle comes to zero as it decelerates, and ultimately reverses. When α is negative, say $= -\beta$, the parameter x_2 cannot increase beyond the value $1/\beta$, and it then decreases again. A finite amount of matter is created and then collapses on itself. Just as each particle is born with the outward velocity of light during the outward expansion, so each particle after being decelerated to rest relative

to the observer is then accelerated inwards and it acquires the inward velocity of light at precisely the same distance from the observer, on its inward return journey, as when it was born It then suffers the converse fate, and disappears from view, in other words it is annihilated The inward collapse is accompanied by the complete annihilation of all the particles created during the outward expansion, until the immediate surroundings of the observer are annihilated

496. It has sometimes been held that these oscillating universes of general relativity provide for cycles of evolution, and for this reason they have been favourably considered Actually they have not been previously analysed in the above detail as regards their appearance to observation, so that it has not been previously realized how remarkably and irretrievably they stand in conflict with experience They imply not only the creation of matter within experience, like the parabolic and hyperbolic systems, they imply also the disappearance of the same matter on the return journey They must be totally and absolutely rejected as irrational They are the fantastic weavings of the mathematical loom, orgies of mathematical licence, divorced from experience They are possible only in the sense that in a dream everything is possible

497. Is the case any better with relativistic systems for which $\lambda \neq 0$, assuming for the sake of argument that nature contains such a constant? If $\lambda < 0$, inward accelerations are increased, and the creation-of-matter effect is but intensified, matter must be created in still greater abundance at the expanding frontier as the newly formed particles lag behind the uniformly expanding frontier with enhanced deceleration If $\lambda > 0$, outward repulsions are introduced These diminish the necessary rate of creation, and we have one case where no creation occurs at all, namely Einstein's original static system, but this end is achieved only by depriving the system of all motion, and no one has yet given a rational account of the world-picture in a static 'spherical' universe, where every particle is in two places all the time, and in any case this contradicts the observed expansion, and sets up an absolute standard of rest Apart from this particular case, as the outward repulsions increase with distance, the Doppler shift increases to infinity, and thereafter the particle concerned disappears from observation, and so is annihilated as far as experience is concerned Positive values of λ only serve to increase the rate of

annihilation These statements could be readily verified by an analysis in terms of observable relations of the formal results of Chapter XVI for $\lambda \neq 0$, using the methods of the present chapter

498. We have now achieved the object of Part IV of this work We have built a bridge joining kinematic cosmology to current relativistic cosmology, by showing how the world-picture of the kinematic models is approximated to under various circumstances by hyperbolic relativistic models This shows that there is an underlying content common to kinematic relativity and to so-called 'general' relativity But the resemblance broke down on further inspection All general relativity models fail because they invoke a creation or annihilation of matter within the experience of the observer, and precisely at this point kinematic relativity avoids this catastrophe, by constructing systems invoking neither creation nor annihilation of matter but already containing, in the experience of any observer, all the particles ever observable, i e wholly observable at any epoch in the experience of any observer

499. Kinematic relativity thus survives as the only system capable of giving a rational account of the world Further, it is not included in 'general' relativity, though it has well-defined relations with it

Necessity of uniform velocities for fundamental particles

500. We can now give a partial answer to a question raised at an earlier stage of this work We began by constructing the kinematics of equivalent observers in arbitrary, possibly accelerated, relative motion, but for the sake of simplicity, and to avoid mathematical difficulties, we applied this kinematics in detail only to *uniform* relative motion The structure we have erected contains as a fundamental element the presence of equivalent observers in uniform relative motion What would have been the course of our analysis if we had carried it out for equivalent accelerated observers?

501. There would have been, as before, a frontier expanding with velocity c But the particles in this frontier would have been subject to acceleration or deceleration, in the experience of the observer at the origin There would thus have been a creation or an annihilation of matter at the expanding frontier creation in the case of deceleration, as the particles lagged behind the light-sphere, in order to fill in the otherwise-arising gap and ensure the continued centrality of each

particle in the field of the rest, annihilation in the case of acceleration, as particles overtook the frontier and passed out of observation In the frontier itself, therefore, the material particles must move with the light, with the constant velocity c in the experience of the observer at the origin (itself a conventional constant introduced as a means of describing experience) But we have seen how deep-lying is the velocity-distance proportionality, how it turns up from the condition of hydrodynamical continuity, from kinematics, or in satisfaction of the cosmological principle, in Newtonian cosmology and in relativistic cosmology Thus on any analysis we may expect the law $v = rF(t)$ for the experience of the observer at the origin, and since v is constant in time for some particles, namely the particles in the expanding light-frontier, by proportional parts v is constant for all fundamental particles, $F(t)$ must be $1/t$ Thus the constancy of relative velocity for the fundamental particles is an essential condition for the avoidance of creation or annihilation of matter Our analysis is thus probably the most general analysis that could be developed for obtaining the solution of the cosmological problem I state the foregoing considerations with reserve, but they would appear to justify our at-first-tentative restriction to uniformly moving equivalent particles

Did the universe originate as a point?

502. A further point which now becomes clear is the following A feature common to the kinematic systems, to the Newtonian systems, and to the 'general' relativity systems with $\lambda = 0$ (in fact to all general relativity systems save those that start from the peculiar Einstein static system) is that at $t = 0$ the systems shrink to a point Is this an accidental mathematical feature, or is it an essential property of the actual universe? Must we believe that the universe actually originated in a point?

The way to answer this question is to consider what would have happened if at $t = 0$ the system had a finite extension, in the experience of an observer O Any system, to satisfy the cosmological principle, must include particles moving with all speeds up to that of light, otherwise there would be a maximum velocity ($< c$) and there would be an accessible boundary But, at any epoch at which the c-moving particles are observed, the local time t' *as observed* is zero—a feature, again, common to kinematic and general relativity systems.

Hence at the most distant, most swiftly moving particles, the local state of affairs must be identical with that in the past experience of the observer at the origin at $t = 0$, i e must be consistent with a finite extension But since these particles are moving with the speed c, they are subject to a complete Lorentz contraction to a point, i e to the observer at the origin, the particles appear infinitely close together, and in the frontier at actually zero separation This does not rule out the possibility of a finite separation in the experience of the particles near the frontier, but I find it difficult to envisage any mathematical description of the state of affairs which provides for a finite separation at $t' = 0$ and $t = 0$ yet permits an apparent infinite closeness of the particles concerned as viewed from a distance If there was a finite separation at $t = 0$, it would be possible to contemplate events for negative values of t, when contraction must have been occurring, and then there should be a distant portion of the universe now visibly contracting, for observation would not now be confined to the interior of the expanding light-sphere Beyond its boundary there would be other particles visible, moving with speeds less than c, condensing towards the place where the velocity equalled c, and so disobeying the velocity-distance proportionality I find it impossible to envisage a violation of the velocity-distance proportionality compatible with the satisfaction of the condition of hydrodynamical continuity—the conservation of particle-number All things considered, it seems impossible to give any rational account of a complete world (satisfying all conditions) enjoying a finite separation at $t = 0$ We must therefore conclude that there was actually a reduction of the world to a point at $t = 0$ This is provided for fully in the kinematic models, both simple and statistical, for the sub-systems have then not yet begun to separate In the kinematic systems the primeval singularity at $t = 0$ in the past history of any observer O is always the exact counterpart of the present singularity (in his experience) now observable at the frontier Our conclusion, then, is that the world was actually created as a point †

503. The mathematical investigations of this chapter show that present-existing relativistic cosmologies are essentially only at the hydrodynamic stage, and hence it is only possible to compare them

† This conclusion was first expressed by Lemaître, in particular at the B A discussion of 1931 (London), but I wish expressly to dissociate myself from any adherence to his view that the world originated as a super-radioactive atom

with our simple kinematic model, as has been done The features of our statistical kinematic model, which have been seen to go so far towards an understanding of the more detailed structural features of the universe, are beyond the present resources of 'general' relativity It is easy to calculate the trajectory of a single free particle in an expanding curved space of general relativity, but to introduce a large number of free particles, statistically distributed, each following a geodesic, would alter the metric (for this is required to be constructed in accordance with the totality of particles present), their mutual gravitation would affect the paths Our method of choosing a space first, free from conditions imposed by the material occupying it, avoids this difficulty

I conclude with a few remarks on 'general' relativity

Remarks on general relativity

504. In our procedure, we make observations first and construct coordinates by combinations of them General relativity begins quite differently If we follow the exposition of Eddington's *Mathematical Theory of Relativity*, we find the following

'Everything connected with location which enters into observational knowledge is contained in a relation of extension between pairs of events This relation is called the interval, and its measure is denoted by *ds* '

But no rule is given for converting observations into a number *ds* It remains at this stage a purely conceptual quantity

It is next supposed that an observer overlays a mesh-system on the external world, and locates an event by four coordinates x, y, z, t

But no rule is given for ascertaining the coordinates of an event from observations of it

We next find

'We have to keep side by side the two methods of describing the configurations of events by coordinates and by the mutual intervals —the first for its conciseness and the second for its immediate absolute significance It is therefore necessary to connect the two modes of description by a formula which will enable us to pass readily from one to the other The particular formula will depend on the coordinates chosen as well as on the absolute properties of the region of the world considered, but it appears that in all cases the formula is included in the following general form ' The author then

proceeds to write down the most general quadratic form in the differentials of coordinates, namely

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu \qquad (\mu, \nu = 1, \ldots, 4) \tag{40}$$

Waiving the fact that neither the numerical value of ds^2 nor the numerical values of the coordinates x^μ have been defined in terms of observations, we have here a fundamental assumption. (We notice that the form is held to depend on the '*absolute* properties of the region of the world considered'. It is curious to begin an account of *relativity* by postulating the existence of undefined 'absolute properties' of a region, when we have previously been told that 'everything we can know about the configuration of events' is contained in a 'relation of extension' and when we notice that all we know of the world are observations of events.) The author admits that this is not the most general case conceivable, that we might have a world in which ds^2 was a general quartic function of the dx's. It might of course be any function whatever, not necessarily homogeneous, why should dt appear in the same way as dx? Actually all these statements are of purely conceptual content, for they make no contact with actual measures.

It is then *assumed* that in a small region (40) is reducible to

$$ds^2 = -c^2dt^2+dx^2+dy^2+dz^2, \tag{40'}$$

and various physical facts are then derived by assuming that a clock-difference measures an interval ds of the type considered. But no immediate justification is given for the unexplained introduction of the minus sign in (40').

Enough has been said to show the entirely conceptual character of the procedure in this presentation. A concept is something which I and the reader do not choose to define, but which, we agree, means something, and the same thing, to both of us. Eddington's introduction to relativity involves a generous use of concepts, but I am personally quite unable to attach any meaning to his phraseology.

505. There is a conceptual element in relativity, but its introduction occurs at a much later stage than appears in Eddington's formulation. From the axiom of the constancy of the velocity of light to observers in uniform relative motion, assuming that velocity has been satisfactorily defined by means of measures with clocks and rigid length-scales, it is shown that if one observer A, from his measures with clocks and rigid length-scales, using Euclidian geometry,

attaches Cartesian coordinates x, y, z, t and $x+dx, y+dy, z+dz, t+dt$ to two given events, and if a second observer B, moving with uniform velocity V relative to A, attaches coordinates x', y', z', t' and $x'+dx', \; , t'+dt'$ to the same events, then

$$c^2dt^2-dx^2-dy^2-dz^2 = c^2dt'^2-dx'^2-dy'^2-dz'^2 \qquad (41)$$

independent of the value of V The common value of these two equal quadratic expressions may now be *defined* to be ds^2, the square of the 'interval' between the events

The striking thing about the formulae

$$ds^2 = c^2\,dt^2-\sum dx^2 = c^2\,dt'^2-\sum dx'^2$$

is not the equality of value of two quadratic functions of the coordinate differentials, but the *conservation of form* of the two quadratic expressions Any observer whatever, provided only he knows himself to be in uniform relative motion with respect to either of A or B, but without using the *value* of his relative velocity, can observe the two events, determine their Cartesian coordinates and epochs, subtract them and then form the summation $c^2\,dt''^2-\sum dx''^2$ with the full assurance that the number he gets is the same number as that got by A or B The rule he has to use for finding what he is to call ds^2 is precisely the same rule as that used by A and B, the actual details of the relations connecting say (x, y, z, t) with (x'', y'', z'', t'') are immaterial—the form is prescribed

It is this factor of conservation of form, applied to the description of systems of moving particles and their accelerations, which we have used for constructing our kinematic relativity The principle has been in all cases *equivalent observers use the same rules* They use the same rules for constructing coordinates out of observations, and when systems have been constructed satisfying the same descriptive rules, the equivalent observers once more use the same rules for calculating accelerations Thus our kinematic relativity is a generalization of the so-called 'restricted' theory of relativity carried out in the spirit of the latter theory

506. Einstein's generalization of 'restricted' relativity proceeds quite differently Einstein's procedure was to assume that given any two events described by conceptual coordinates derived from the superposition of a conceptual mesh-system on events, then the *value* of ds^2 was conserved for all transformations of coordinates Conservation of form, the most striking feature of the consequences of experi-

ments with light-signals, was sacrificed in favour of conservation of value Relation (40′) was then held to be replaceable in general by (40) This is sometimes expressed by saying that the world of events is supposed describable, in 'general' relativity, by a Riemannian space The statement is, however, meaningless until we have provided some rules for passing from coordinates to observations or vice versa

507. This rule is simply that light-rays are supposed to follow nul geodesics in the Riemannian space $ds = 0$ along a path $\delta \int ds = 0$ This rule provides a complete scheme of transformation from coordinates (the coordinates of any specified event), in any chosen metric (formula of the type (40) for ds^2), to the observations that could be made by any specified observer (also specified by coordinates) who observed that event But it does not give any rule for passing from observations of events to the associated metric In fact it by no means follows from the principles of general relativity that any set of observations of any set of events can be described by a metric of the type (40), and our kinematic systems are examples of systems, duly defined by observations, which cannot be so described

508. This is the optical content of the 'general' theory of relativity Its kinematic content is now given by the further assumption that a free test-particle will follow a geodesic in the Riemannian space, $\delta \int ds = 0$

509 But the structure so far obtained is still meaningless We choose a metric, and we predict from this that a free test-particle, liberated in such and such a way, will pursue such and such an observable path But we do not know what we mean by having chosen a metric, in terms of observation We do not yet know what material occupies the continuum so defined In the presence of what other particles does the test-particle move in the way predicted? This question is answered by supplementing the conceptual scheme by 'field equations'. These state that if from the metric we calculate the tensor $G_{\mu\nu}+g_{\mu\nu}(\lambda-\frac{1}{2}G)$, then the matter-in-motion present is defined by a second tensor $-\kappa T_{\mu\nu}$ equal to the tensor just calculated, where the components of $T_{\mu\nu}$ define the motion, pressure, and density of the matter present in terms of the conceptual coordinate mesh-system already used to locate events It is a self-consistent property of the

whole formulation that the motions defined by $T_{\mu\nu}$ obey the stationary principle $\delta \int ds = 0$

510. Until the field equations are stated, we do not know the matter present, but we do know the paths of free particles in the conceptual scheme We can regard the question in the following way Let a metric be chosen, and the geodesics $\delta \int ds = 0$ calculated Let constraints be introduced so as to compel any given test particle to follow a geodesic Then introduce *matter* into the system until the required constraints reduce to zero Then this is the matter implied to be present This method of regarding the situation exhibits the analogy between the kinematic content of the 'general' theory of relativity and our kinematic relativity For we first constrained our particles present to follow prescribed paths, then showed that free particles would follow the same paths without constraints

511 It is to be noted that the procedure can in principle be reversed, we can state $T_{\mu\nu}$, and then derive the $g_{\mu\nu}$'s corresponding to it But a coordinate mesh-system must be chosen before $T_{\mu\nu}$ can be described, and this is usually stated in conceptual terms, not in terms of the observations that would in fact determine $T_{\mu\nu}$ If $T_{\mu\nu}$ be stated, the process can be described as one of building up a space appropriate to the matter-in-motion present according to certain rules, namely the rules contained in the field equations

512 The formal relationships of the theory thus constructed are satisfied whatever value λ has Thus if the $g_{\mu\nu}$'s are given, the matter present is in fact indeterminate, since the principles of the theory do not fix the value of λ But the geodesics remain determinate and independent of λ General relativity offers in fact a wide choice of distributions of matter-in-motion all compatible with the same set of free trajectories

513. In its applications to the cosmological problem, general relativity follows the first-mentioned procedure It first chooses a metric, partly on grounds of symmetry, partly from considerations of 'cosmic time', this is equivalent to choosing a space and thus implying a set of trajectories that would be followed by free particles The field equations are then employed to ascertain (with the ambiguity consequent on the admission of the cosmical constant λ) the matter-in-

motion implied to be present Thus the space is first chosen, and then filled with matter according to the 'laws of nature' which are assumed as part of the theory Lastly the whole structure is convertible into the observations that would actually reveal it, by means of the light-paths defined by $ds = 0$, $\delta \int ds = 0$

514. The above account is sufficient to disclose the essential differences between the procedure of general relativity and our kinematic procedure As Eddington has written, in general relativity 'we start with a general theory of world-structure and work down to the experimental consequences' The general theory is stated first in terms of concepts and a conceptual symbolism, and the laws of the structure are stated in the same conceptual terms In the kinematic treatment, too, we start by choosing a space—the ordinary space of physics for each observer We relate different observers' different spaces But we begin with no theory of world-structure, or the assumed existence of 'laws of nature' We begin with the observations that could in principle be made, and work upwards to the regularities that these observations, to be compatible, must satisfy Whilst in certain cases general relativity chooses a finite space and then in accordance with its own laws fills it necessarily with matter of a finite density, thus arriving at a finite number of particles in the universe, our method leaves the occupation of the space to be determined by our principle of selection of the systems to be considered, namely the cosmological principle We make no *a priori* decision as to whether the world contains or does not contain an infinite number of particles, or as to whether it occupies a finite or an infinite portion of the observer's space as customarily used in physics The procedure settles these questions for itself, and leads to the conclusion that the world contains an infinite number of particles distributed inside a finite volume of the above space It automatically leads us to consider the properties of 'open' sets of points, which play no part in general relativity The procedure generates descriptions of sequences of events which may be called 'laws of nature', but they are not inferred as consequences of obedience to supposed laws *governing* nature The structure is thus completely free from any appeal to specific theories about nature As originally maintained by Eddington, but on different grounds, we conclude that as far as gravitation is concerned there is no such thing as 'a law of gravitation', only

a totality of possible gravitational situations, of which we have here described in detail those satisfying the cosmological principle Gravitation disappears either as an assumed law or as a consequence of supposed properties of a supposed entity 'space-time', sets of compatible accelerations of particles, in one another's presence, alone survive

Conscious of many imperfections, and of hosts of fascinating problems awaiting further investigation, we now leave the subject

APPENDIX

MATHEMATICAL NOTES

NOTE 1 *The world-picture* (§ 77)

WE have established from kinematic considerations the formulae

$$n(t) = N\exp\left(-3\int_T^t F(t)\,dt\right)$$

and
$$\frac{1}{r}\frac{dr}{dt} = F(t),$$

where d/dt denotes differentiation following the motion

At time of observation t, the number of particles or nebulae actually within the range $(r, r+dr)$ and inside a solid angle $d\omega$ is

$$r^2\,drd\omega\,n(t)$$
$$= N\exp\left(-3\int_T^t F(t)\,dt\right)r^2\,drd\omega$$

In the photograph, this same number of nebulae is visible at observed distances between r_1 and r_1+dr_1, where

$$t_1 = t-r_1/c,$$
$$r = r_1\exp\left(\int_{t_1}^t F(t)\,dt\right)$$
$$= r_1\exp\left(\int_{t-r_1/c}^t F(t)\,dt\right),$$

and accordingly, keeping t constant in differentiating,

$$dr = dr_1\exp\left(\int_{t-r_1/c}^t F(t)\,dt\right)\left[1+\frac{r_1}{c}F\left(t-\frac{r_1}{c}\right)\right]$$

Hence in terms of r_1 and dr_1 the above number of nebulae is equal to

$$N\exp\left(-3\int_T^t F(t)\,dt\right)\left[1+\frac{r_1}{c}F\left(t-\frac{r_1}{c}\right)\right]\exp\left(3\int_{t-r_1/c}^t F(t)\,dt\right)r_1^2\,dr_1\,d\omega$$

This is equal to

$$n(t)\left[1+\frac{r_1}{c}F\left(t-\frac{r_1}{c}\right)\right]\exp\left(3\int_{t-r_1/c}^t F(t)\,dt\right)r_1^2\,dr_1\,d\omega,$$

or approximately, if the variation of $F(t)$ in an interval r_1/c is sufficiently small,

$$n(t)\left[1+\frac{r_1}{c}F\left(t-\frac{r_1}{c}\right)\right]\left[1+3\frac{r_1}{c}F\left(t-\frac{r_1}{c}\right)\right]r_1^2\,dr_1\,d\omega$$

But
$$\frac{r_1}{c}F\left(t-\frac{r_1}{c}\right) = \frac{r_1}{c}F(t_1) = \frac{v_1}{c},$$

where v_1 is the observed velocity of the nebula photographed as at distance r_1 Hence the number of nebulae between photographed distances r_1 and r_1+dr_1 at time of observation t is

$$n(t)\left(1+4\frac{v_1}{c}\right)r_1^2\,dr_1\,d\omega$$

If this is written as $\nu(r_1, t)r_1^2\,dr_1\,d\omega$ we have

$$\nu(r_1,t) = n(t)\left(1+4\frac{v_1}{c}\right)$$

This is the formula quoted

NOTE 2 *Solution of the functional equations for the invariant velocity-distribution* (§ 89)

The functional equations in question are

$$f(u,v,w) = f\left(\frac{u-U}{1-uU/c^2}, \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}\right)\frac{(1-U^2/c^2)^2}{(1-uU/c^2)^4},$$

and two similar equations To solve them, make the substitution

$$f(u,v,w) = \frac{F(u,v,w)}{\left(1-\dfrac{u^2+v^2+w^2}{c^2}\right)^2}$$

We have

$$1-\frac{1}{c^2}\left[\left(\frac{u-U}{1-uU/c^2}\right)^2+\left(\frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}\right)^2+\left(\frac{w(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}\right)^2\right]$$
$$\equiv \frac{1-U^2/c^2}{(1-uU/c^2)^2}\left[1-\frac{u^2+v^2+w^2}{c^2}\right]$$

Hence the functional equation becomes

$$F(u,v,w) = F\left(\frac{u-U}{1-uU/c^2}, \frac{v(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}, \frac{w(1-uU/c^2)^{\frac{1}{2}}}{1-uU/c^2}\right)$$

Without appealing to the theory of invariants under the Einstein velocity-transformation, we may solve this functional equation in an elementary way as follows Since the equation is to hold good for all U, it is certainly satisfied when U is small Expanding the right-hand side by Taylor's theorem to the first power of U we find

$$-\left(1-\frac{u^2}{c^2}\right)\frac{\partial F}{\partial u}+\frac{uv}{c^2}\frac{\partial F}{\partial v}+\frac{uw}{c^2}\frac{\partial F}{\partial w} = 0$$

This may be written in the form

$$\frac{1}{u}\frac{\partial F}{\partial u} = \frac{1}{c^2}\left(u\frac{\partial F}{\partial u}+v\frac{\partial F}{\partial v}+w\frac{\partial F}{\partial w}\right),$$

whence by symmetry the other two equations must give

$$\frac{1}{u}\frac{\partial F}{\partial u} = \frac{1}{v}\frac{\partial F}{\partial v} = \frac{1}{w}\frac{\partial F}{\partial w} = \frac{1}{c^2}\left(u\frac{\partial F}{\partial u}+v\frac{\partial F}{\partial v}+w\frac{\partial F}{\partial w}\right)$$

Put $\qquad u^2 = \alpha, \qquad v^2 = \beta, \qquad w^2 = \gamma$

Then the equations become

$$\frac{\partial F}{\partial \alpha} = \frac{\partial F}{\alpha\beta} = \frac{\partial F}{\partial \gamma} = \frac{1}{c^2}\left(\alpha\frac{\partial F}{\partial \alpha}+\beta\frac{\partial F}{\partial \beta}+\gamma\frac{\partial F}{\partial \gamma}\right)$$

The solution of the first two of these is of the form

$$F \equiv \Phi(\alpha+\beta+\gamma),$$

as is immediately proved The remaining relation then gives

$$\Phi'(\alpha+\beta+\gamma) = \frac{\alpha+\beta+\gamma}{c^2}\Phi'(\alpha+\beta+\gamma),$$

whence $$\Phi' \equiv 0, \qquad \Phi \equiv \text{const},$$

so that $$F \equiv \text{const}$$

The solution of the set of functional equations must accordingly be of the form

$$f(u,v,w) \equiv \frac{B}{c^3\left(1-\dfrac{u^2+v^2+w^2}{c^2}\right)^2},$$

where B is any constant It is now readily verified by direct substitution that this is an exact solution for any U

NOTE 3 *Solution of the functional equations for the acceleration* (§ 101)

The equations in question are (24), § 101, together with the similar equations obtained by replacing $(U,0,0)$ by $(0,U,0)$ and $(0,0,U)$, $\mathbf{P}'$, $\mathbf{V}'$, and t' being connected with $\mathbf{P}$, $\mathbf{V}$, t by the Lorentz transformation The equations must be identities in the seven variables x,y,z,t,u,v,w

To determine the functions f, g, h, substitute

$$f(\mathbf{P},\mathbf{V},t) \equiv Y\alpha(\mathbf{P},\mathbf{V},t)$$
$$g(\mathbf{P},\mathbf{V},t) \equiv Y\beta(\mathbf{P},\mathbf{V},t)$$
$$h(\mathbf{P},\mathbf{V},t) \equiv Y\gamma(\mathbf{P},\mathbf{V},t),$$

where Y is a function of $\mathbf{V}$ given by (25), § 101 Then $f(\mathbf{P}',\mathbf{V}',t') \equiv Y'\alpha(\mathbf{P}',\mathbf{V}',t')$, etc Inserting these in the given equations (24) and using (27′) we find the functional equations

$$\alpha(\mathbf{P}',\mathbf{V}',t') = \frac{(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}\alpha(\mathbf{P},\mathbf{V},t),$$

$$\beta(\mathbf{P}',\mathbf{V}',t') = \beta(\mathbf{P},\mathbf{V},t)+\frac{vU/c^2}{1-uU/c^2}\alpha(\mathbf{P},\mathbf{V},t),$$

$$\gamma(\mathbf{P}',\mathbf{V}',t') = \gamma(\mathbf{P},\mathbf{V},t)+\frac{wU/c^2}{1-uU/c^2}\alpha(\mathbf{P},\mathbf{V},t)$$

Next, make the substitution

$$\alpha(\mathbf{P},\mathbf{V},t) \equiv (x-ut)L(\mathbf{P},\mathbf{V},t),$$
$$\beta(\mathbf{P},\mathbf{V},t) \equiv (y-vt)M(\mathbf{P},\mathbf{V},t),$$
$$\gamma(\mathbf{P},\mathbf{V},t) \equiv (z-wt)N(\mathbf{P},\mathbf{V},t)$$

Then $\alpha(\mathbf{P}',\mathbf{V}',t') = (x'-u't')L(\mathbf{P}',\mathbf{V}',t')$, etc By use of the Lorentz formulae and the velocity-addition formulae we find that

$$x'-u't' = \frac{(1-U^2/c^2)^{\frac{1}{2}}}{1-uU/c^2}(x-ut),$$

$$y'-v't' = \frac{(y-vt)+(U/c^2)(vx-uy)}{1-uU/c^2},$$

$$z'-w't' = \frac{(z-wt)+(U/c^2)(wx-uz)}{1-uU/c^2}$$

The functional equations for α, β, γ then become

$$L(\mathbf{P}', \mathbf{V}', t') = L(\mathbf{P}, \mathbf{V}, t),$$

$$[(y-vt)+(U/c^2)(vx-uy)]M(\mathbf{P}', \mathbf{V}', t') = (y-vt)(1-uU/c^2)M(\mathbf{P}, \mathbf{V}, t)+ \\ +(vU/c^2)(x-ut)L(\mathbf{P}, \mathbf{V}, t),$$

$$[(z-wt)+(U/c^2)(wx-uz)]N(\mathbf{P}', \mathbf{V}', t') = (z-wt)(1-uU/c^2)N(\mathbf{P}, \mathbf{V}, t)+ \\ +(wU/c^2)(x-ut)L(\mathbf{P}, \mathbf{V}, t)$$

The first of these equations shows that L is an invariant function of x, y, z, t, u, v, w under a Lorentz transformation It is easily proved that the only invariants under such transformations are

$$X \text{ and } Z^2/Y,$$

where X, Y, Z are given by (25) and satisfy (27), (27′), (27″) Hence

$$L \equiv L(X, Z^2/Y)$$

The analogous equations obtained by consideration of $(0, U, 0)$ and $(0, 0, U)$ show that

$$M(\mathbf{P}', \mathbf{V}', t') = M(\mathbf{P}, \mathbf{V}, t),$$

$$N(\mathbf{P}', \mathbf{V}', t) = N(\mathbf{P}, \mathbf{V}, t)$$

These may be written accordingly as $M(X, Z^2/Y)$, $N(X, Z^2/Y)$

The functional equations for M and N now simplify, on noticing that

$$[y-vt+(U/c^2)(vx-uy)]-(y-vt)(1-uU/c^2) \equiv (vU/c^2)(x-ut),$$

$$[z-wt+(U/c^2)(wx-uz)]-(z-wt)(1-uU/c^2) \equiv (wU/c^2)(x-ut),$$

to

$$M(X, Z^2/Y) \equiv L(X, Z^2/Y),$$

$$N(X, Z^2/Y) \equiv L(X, Z^2/Y)$$

Accordingly, the solution of the original set of functional equations is of the form

$$f = Y(x-ut)G(X, Z^2/Y), \qquad g = Y(y-vt)G(X, Z^2/Y),$$

$$h = Y(z-wt)G(X, Z^2/Y)$$

We now notice that f, g, h are of dimensions (length)(time)$^{-2}$ Consequently $G(X, Z^2/Y)$ is of the form $X^{-1}G_1(X, Z^2/Y)$, where G_1 is dimensionless Since X and Z^2/Y are of the dimensions (time)^2, G_1 must be a function of the single dimensionless variable Z^2/XY or ξ, and we may write it $G_1 \equiv G(\xi)$ The solution is accordingly

$$f = (x-ut)\frac{Y}{X}G(\xi),$$

$$g = (y-vt)\frac{Y}{X}G(\xi),$$

$$h = (z-wt)\frac{Y}{X}G(\xi),$$

or, in vector form,

$$\mathbf{g} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi)$$

The substitutions we have employed obtain the most general solution of the original equations They are, however, by no means obvious, and they were determined only after the solution had first been found by more tentative methods The method employed in my original paper† was to express the vector $\mathbf{g}$ in the form

$$\mathbf{g} = A\mathbf{P}+B\mathbf{V}+C(\mathbf{P}\wedge\mathbf{V}),$$

† *Zeits fur Astrophys* **6**, 53, 1933

where A, B, C are three scalar functions of x, , t, , w to be determined The condition that this expression shall transform as an acceleration was then worked out, using easy four-dimensional vector calculus It appeared that A/Y must be an invariant, that $(At+B)/Y^{\frac{3}{2}}$ must also be an invariant, and finally that $At+B$ must be zero Thus $B = -At$, and the result follows The appearance of the symbol t in the apparently odd form $At+B$ arose from expressing the scalar product $\mathbf{P}\ \mathbf{V}$ in the form $c^2(t-Z)$

NOTE 4 *The vector integrals* (§ 168)

We have from the equations of motion

$$\frac{d\mathbf{P}}{dt} = \mathbf{V}, \qquad \frac{d\mathbf{V}}{dt} = (\mathbf{P}-\mathbf{V}t)\frac{Y}{X}G(\xi),$$

that

$$\frac{d}{dt}(\mathbf{P}-\mathbf{V}t) = -t\frac{d\mathbf{V}}{dt} = -(\mathbf{P}-\mathbf{V}t)t\frac{Y}{X}G(\xi),$$

$$\frac{d}{dt}(\mathbf{P}\wedge\mathbf{V}) = \mathbf{P}\wedge\frac{d\mathbf{V}}{dt} = -(\mathbf{P}\wedge\mathbf{V})t\frac{Y}{X}G(\xi)$$

Also

$$\begin{aligned}\frac{d}{dt}(Z^2-XY) &= 2Z\frac{dZ}{dt}-\frac{dX}{dt}Y-X\frac{dY}{dt}\\ &= 2Z\Big[Y\Big\{1+G(\xi)\frac{X-tZ}{X}\Big\}\Big]-2YZ+2Y(tY-Z)G(\xi)\\ &= -2G(\xi)\frac{Y}{X}t(Z^2-XY),\end{aligned}$$

whence

$$\frac{d}{dt}(Z^2-XY)^{\frac{1}{2}} = -t\frac{Y}{X}G(\xi)(Z^2-XY)^{\frac{1}{2}}$$

To establish identity (17) of § 167, we have

$$\begin{aligned}\theta^2 &\equiv \Big(t-\frac{\mathbf{P}\ \mathbf{V}}{c^2}\Big)^2-\Big(t^2-\frac{\mathbf{P}^2}{c^2}\Big)\Big(1-\frac{\mathbf{V}^2}{c^2}\Big)\\ &= \frac{\mathbf{V}^2t^2}{c^2}-2\frac{(\mathbf{P}\ \mathbf{V})}{c^2}t+\frac{\mathbf{P}^2}{c^2}-\frac{1}{c^4}[\mathbf{P}^2\mathbf{V}^2-(\mathbf{P}\ \mathbf{V})^2]\\ &= \frac{1}{c^2}(\mathbf{P}-\mathbf{V}t)^2-\frac{1}{c^4}(\mathbf{P}\wedge\mathbf{V})^2\end{aligned}$$

To solve equations (13) and (14) of § 166, insert for $\mathbf{P}$ from (13) in (14) We obtain

$$(\mathbf{f}\theta+\mathbf{V}t)\wedge\mathbf{V} = c\mathbf{l}\theta,$$

or, since $\mathbf{V}\wedge\mathbf{V} = 0$,

$$(\mathbf{f}\wedge\mathbf{V}) = c\mathbf{l}$$

Multiply this vectorially by $\mathbf{f}$ Then by the continued vector product theorem

$$-\mathbf{f}(\mathbf{V}\ \mathbf{f})+\mathbf{V}f^2 = c\mathbf{l}\wedge\mathbf{f}$$

But the equation $\mathbf{f}\wedge\mathbf{V} = c\mathbf{l}$ is impotent to determine the component of $\mathbf{V}$ along $\mathbf{f}$ Hence

$$\mathbf{V} = c\frac{\mathbf{l}\wedge\mathbf{f}}{\mathbf{f}^2}+\lambda\mathbf{f},$$

where λ is undetermined We have then

$$\begin{aligned}\mathbf{P} &= \mathbf{f}\theta+\mathbf{V}t\\ &= c\frac{\mathbf{l}\wedge\mathbf{f}}{\mathbf{f}^2}t+(\lambda t+\theta)\mathbf{f}\end{aligned}$$

NOTE 5 *Derivation of Boltzmann's equation* (§ 209)

If x, y, z are the components of $\mathbf{P}$, x_1, y_1, z_1 of $\mathbf{P}_1$, u, v, w of $\mathbf{V}$, u_1, v_1, w_1 of $\mathbf{V}_1$, g_1, g_2, g_3 of $\mathbf{g}$, then

$$\frac{do_1\,d\tau_1}{do\,d\tau} = \frac{\partial(x_1, y_1, z_1, u_1, v_1, w_1)}{\partial(x, y, z, u, v, w)},$$

where

$$x_1 = x+u\,\Delta t, \qquad y_1 = y+v\,\Delta t, \qquad z_1 = z+w\,\Delta t,$$

$$u_1 = u+g_1\,\Delta t, \qquad v_1 = v+g_2\,\Delta t, \qquad w_1 = w+g_3\,\Delta t$$

Here the Jacobian is

$$\begin{vmatrix} 1, & 0, & 0, & \Delta t, & 0, & 0 \\ 0, & 1, & 0, & 0, & \Delta t, & 0 \\ 0, & 0, & 1, & 0, & 0, & \Delta t \\ \frac{\partial g_1}{\partial x}\Delta t, & \frac{\partial g_1}{\partial y}\Delta t, & \frac{\partial g_1}{\partial z}\Delta t, & 1+\frac{\partial g_1}{\partial u}\Delta t, & \frac{\partial g_1}{\partial v}\Delta t, & \frac{\partial g_1}{\partial w}\Delta t \\ \frac{\partial g_2}{\partial x}\Delta t, & \frac{\partial g_2}{\partial y}\Delta t, & \frac{\partial g_2}{\partial z}\Delta t, & \frac{\partial g_2}{\partial u}\Delta t, & 1+\frac{\partial g_2}{\partial v}\Delta t, & \frac{\partial g_2}{\partial w}\Delta t \\ \frac{\partial g_3}{\partial x}\Delta t, & \frac{\partial g_3}{\partial y}\Delta t, & \frac{\partial g_3}{\partial z}\Delta t, & \frac{\partial g_3}{\partial u}\Delta t, & \frac{\partial g_3}{\partial v}\Delta t, & 1+\frac{\partial g_3}{\partial w}\Delta t \end{vmatrix},$$

which to the first power in Δt expands as

$$1+\Delta t\left(\frac{\partial g_1}{\partial u}+\frac{\partial g_2}{\partial v}+\frac{\partial g_3}{\partial w}\right)$$

The condition that particles are neither created nor destroyed is accordingly given by

$$f(\mathbf{P}+\mathbf{V}\,\Delta t, t+\Delta t, \mathbf{V}+\mathbf{g}\,\Delta t)\left[1+\Delta t\frac{\partial}{\partial \mathbf{V}}\,\mathbf{g}\right] = f(\mathbf{P}, t, \mathbf{V})$$

By the definition of vector derivatives, this requires

$$\frac{\partial f}{\partial t}+\mathbf{V}\,\frac{\partial f}{\partial \mathbf{P}}+\mathbf{g}\cdot\frac{\partial f}{\partial \mathbf{V}}+f\left(\frac{\partial}{\partial \mathbf{V}}\,\mathbf{g}\right) = 0$$

This is the desired equation

NOTE 6 *A velocity relationship* (§ 316)

The problem is to eliminate the unit vector $\mathbf{i}$ from the relation

$$\mathbf{V}_0' = \mathbf{V}_0 - c\mathbf{i}(1-\mathbf{V}_0^2/c^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)}, \tag{1}$$

where

$$\sinh\epsilon = \frac{(\mathbf{V}_0\,\mathbf{i})/c}{(1-\mathbf{V}_0^2/c^2)^{\frac{1}{2}}}. \tag{2}$$

To effect the elimination, multiply (1) vectorially by $\mathbf{V}_0$. Then

$$\mathbf{V}_0' \wedge \mathbf{V}_0 = -c(\mathbf{i}\wedge\mathbf{V}_0)(1-\mathbf{V}_0^2/c^2)^{\frac{1}{2}}\frac{\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)}$$

Taking the square of the modulus of each side we have, since

$$(\mathbf{l} \wedge \mathbf{V}_0)^2 = \mathbf{V}_0^2-(\mathbf{l}\ \mathbf{V}_0)^2,$$

$$(\mathbf{V}_0' \wedge \mathbf{V}_0)^2 = c^2(1-V_0^2/c^2)[V_0^2-c^2(1-V_0^2/c^2)\sinh^2\epsilon]\,\frac{\sinh^2 2\eta_l}{\cosh^2(\epsilon-2\eta_l)}$$

$$= c^2(1-V_0^2/c^2)[V_0^2\cosh^2\epsilon-c^2\sinh^2\epsilon]\,\frac{\sinh^2 2\eta_l}{\cosh^2(\epsilon-2\eta_l)}$$

Also from (1), $$(\mathbf{V}_0'-\mathbf{V}_0)^2 = c^2(1-V_0^2/c^2)\,\frac{\sinh^2 2\eta_l}{\cosh^2(\epsilon-2\eta_l)}$$

Hence

$$c^2(\mathbf{V}_0'-\mathbf{V}_0)^2-(\mathbf{V}_0' \wedge \mathbf{V}_0)^2 = c^2(1-V_0^2/c^2)\,\frac{\sinh^2 2\eta_l}{\cosh^2(\epsilon-2\eta_l)}\,(c^2-V_0^2)\cosh^2\epsilon \tag{3}$$

But multiplying (1) scalarly by $\mathbf{V}_0$,

$$\mathbf{V}_0'\ \mathbf{V}_0 = V_0^2-c^2(1-V_0^2/c^2)\,\frac{\sinh\epsilon\sinh 2\eta_l}{\cosh(\epsilon-2\eta_l)},$$

or $$c^2-\mathbf{V}_0'\ \mathbf{V}_0 = (c^2-V_0^2)\,\frac{\cosh\epsilon\cosh 2\eta_l}{\cosh(\epsilon-2\eta_l)} \tag{4}$$

By (3) and (4), $$\tanh^2 2\eta_l = \frac{c^2(\mathbf{V}_0'-\mathbf{V}_0)^2-(\mathbf{V}_0' \wedge \mathbf{V}_0)^2}{(c^2-\mathbf{V}_0'\ \mathbf{V}_0)^2},$$

which is the desired eliminant

NOTE 7 *The apparent brightness of a receding nebula* (§§ 140, 149)

In order to show that the simple kinematic model gives a continuous back ground of finite intensity, it is necessary to have a formula for the apparent brightness of a receding object as measured photometrically or spectro photometrically by a given observer at a given epoch of his experience The luminosity of an object such as a nebula may be supposed given as a function of local time, the same function for all nebulae on the average. This function, which depends on the evolutionary optical history of the nebula, is entirely unknown, but for our purpose it is sufficient to leave it as an undetermined function, the form of the formula for the apparent brightness will then show sufficiently well whether the background intensity converges to a finite limit in spite of the infinite number of luminous objects present The effect of the recession, of course, is to diminish the apparent brightness, for the recession dilutes the stream of photons in space as measured by the observer This diminution in apparent brightness amounts to complete invisibility as the speed of recession approaches that of light, and hence it comes about that the background intensity, due to the infinity of fast-moving objects with velocities in excess of any given velocity, can be finite

An approximate formula for the effect, valid for small velocities, has been given by Hubble,† from physical considerations But for our purpose of applying the result to velocities approaching that of light, an exact formula is required The following is a revised version of the original investigation given by the author,‡ which has been confirmed by another method by A G Walker ||

† *Astrophysical Journal*, **74**, 71, 1931 ‡ *Zeits fur Astrophys*, **6**, 90, 1933
|| *M N, R A S*, **94**, 162, 1934

As brightness is measured in energy units, it is necessary to depart somewhat from a purely kinematic treatment

Our general method is to begin by considering a nebula at rest and an energy-receiver in motion relative to it We then transform the intensity at the receiver, as measured by the particle-observer on the nebula, into the intensity measured by the particle observer at the receiver The optical properties of the nebula are supposed given, however, in terms of the units appropriate to an observer on the nebula, as this gives the standard behaviour of the nebula We have of course to express the result in terms of the epoch at which the observation in made, in the experience of the observer at the receiver

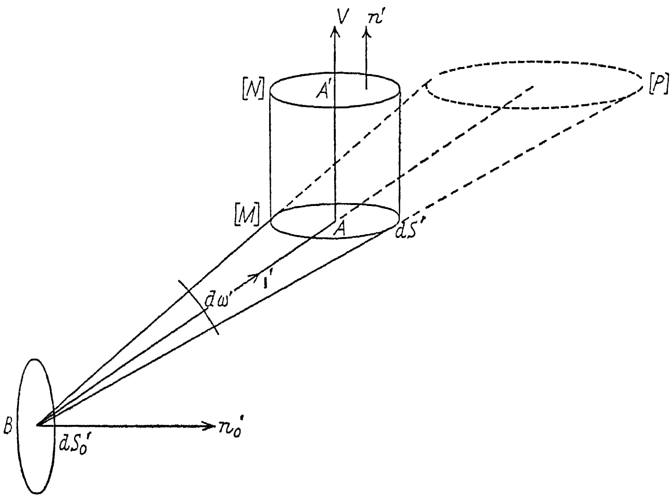


FIG 20 Calculation of radiation from a receding surface element

For convenience we consider first the case of any uniform relative motion of source and receiver, not necessarily one of pure recession Primed symbols refer to B's measures Let dS_0' be a radiating surface centred at the point B, with its normal in the direction of a unit vector $\mathbf{n}_0'$ Let dS' be a receiving surface whose centre, at the epoch t' in B's experience (i e at time t' by B's clock) is at the point A Let $\mathbf{n}'$ be a unit vector normal to dS', and let $\mathbf{V}$ be its velocity Let $BA = r'$, and let $\mathbf{1}'$ be a unit vector in the direction BA, $d\omega'$ the solid angle subtended by dS' at B when dS' is at A Then

$$d\omega' = \frac{dS'(\mathbf{1}'\ \mathbf{n}')}{r'^2}$$

Consider the energy that would be received by dS' from dS_0' in the interval $(t', t'+dt')$ of B's experience If dS' were stationary at A, the amount of energy it would receive during $(t', t'+dt')$ would be, by the general theory of radiation,

$$(I'_{\nu'}\,d\nu')\,dS_0'(\mathbf{i}'\ \mathbf{n}_0')\,d\omega' dt' = \frac{(I'_{\nu'}\,d\nu')\,dS_0'(\mathbf{1}'\ \mathbf{n}_0')\,dS'(\mathbf{i}'\ \mathbf{n}')\,dt'}{r'^2} \tag{1}$$

Here $I'_{\nu'}$ is the specific intensity of radiation from dS_0' in the direction $\mathbf{1}'$, in frequency ν', as measured by the particle observer B When dS' is in motion, the

amount of energy it receives is altered Let $[M]$ denote the position of dS' at t', $[N]$ its position at $t'+dt'$ Project dS' at A conically from B on to the plane of $[N]$, and call this projected position $[P]$ Then the radiation crossing dS' as it passes from $[M]$ to $[N]$, to the first power of the differentials concerned, is equal in amount to the radiation which would have crossed dS' had it passed from $[M]$ to $[P]$ The error is in fact of the order of the energy passing through the curved surface of the cylinder formed by the motion of dS' from $[M]$ to $[N]$, this energy is equal to the difference between the energies that would cross, in the same time, the area dS' stationary in the positions $[M]$ and $[N]$ respectively, which is of a higher order than the differential we are calculating.

Now the amount of energy falling on dS' in its motion from $[M]$ to $[P]$ is less than the amount that would fall on dS' if stationary at $[M]$ by the amount of energy trapped in the truncated cone cut off between $[M]$ and $[P]$ as subtended from B The amount of this energy is equal to the amount of energy falling on a stationary dS' at $[M]$ in one second, namely

$$\frac{(I'_{\nu'}\,d\nu')\,dS'_0(\mathbf{i}'\;\mathbf{n}'_0)\,dS'(\mathbf{i}'\;\mathbf{n}')}{r'^2},$$

multiplied by the time taken by radiation to move from $[M]$ to $[P]$ along BA produced, namely

$$\frac{1}{c}\,\frac{(\mathbf{V}\;\mathbf{n}')\,dt'}{\mathbf{i}'\;\mathbf{n}'}.$$

The product of these expressions is

$$(I'_{\nu'}\,d\nu')\,dS'_0(\mathbf{i}'\;\mathbf{n}'_0)\,\frac{dS'}{r'^2}\left(\frac{\mathbf{V}\;\mathbf{n}'}{c}\right)dt' \tag{2}$$

Subtracting this from (1) we have for the amount of energy passing through the moving dS' in time dt'

$$(I'_{\nu'}\,d\nu')\,dS'_0(\mathbf{i}'\;\mathbf{n}'_0)\,\frac{dS'}{r'^2}\,dt'\left[(\mathbf{i}'\;\mathbf{n}')-\frac{\mathbf{V}\;\mathbf{n}'}{c}\right], \tag{3}$$

as measured by B Call this E'

Take now $\mathbf{V}$ parallel to $\mathbf{i}'$, so that $\mathbf{V} = V\mathbf{i}'$ Then

$$E' = (I'_{\nu'}\,d\nu')\,dS'_0(\mathbf{i}'\;\mathbf{n}'_0)\,\frac{dS'}{r'^2}\,dt'(\mathbf{i}'\;\mathbf{n}')\left(1-\frac{V}{c}\right) \tag{4}$$

To calculate A's measure of this energy, it is sufficient to notice that A receives, in the interval $(t', t'+dt')$ of B's reckoning, $E'/h\nu'$ quanta, each of which he will count as a quantum of energy $h\nu$, where

$$\nu = \nu'\left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}} \tag{5}$$

Thus the energy received by A, in A's measure, is

$$E = E'\frac{h\nu}{h\nu'} = E'\left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}, \tag{6}$$

i e by (4)

$$E = (I'_{\nu'}\,d\nu')\frac{dS'_0(\mathbf{i}'\;\mathbf{n}'_0)\,dS'(\mathbf{i}'\;\mathbf{n}')\,dt'}{r'^2}\;\frac{(1-V/c)^2}{(1-V^2/c^2)^{\frac{1}{2}}} \tag{7}$$

This number must now be expressed in terms of measures used by A The event of the arrival of the light at A has the coordinates $(-r', t')$ in B's experi-

ence (taking for convenience an x-axis from A to B) Also here, $r' = Vt'$, by the recession law and the definition of the origin of t' The formulae

$$x' = \frac{x - Vt}{(1-V^2/c^2)^{\frac{1}{2}}}, \qquad t' = \frac{t - Vx/c^2}{(1-V^2/c^2)^{\frac{1}{2}}},$$

give accordingly here

$$-Vt' = \frac{0 - Vt}{(1-V^2/c^2)^{\frac{1}{2}}}, \qquad t' = \frac{t}{(1-V^2/c^2)^{\frac{1}{2}}},$$

where $(0, t)$ is the description of the event by A Hence

$$r' = \frac{Vt}{(1-V^2/c^2)^{\frac{1}{2}}}, \tag{8}$$

$$dt' = \frac{dt}{(1-V^2/c^2)^{\frac{1}{2}}} \tag{9}$$

Further, since dS and dS_0 are directly receding from one another, and $\mathbf{i}'$ is the direction of mutual recession in B's reckoning, we have $\mathbf{i}' = \mathbf{i}$, and $dS'_0(\mathbf{i}'\ \mathbf{n}'_0)$ and $dS'(\mathbf{i}'\ \mathbf{n}')$, being projections of areas normal to the direction of motion, are unaltered by transformation from B's measures to A's, and so

$$dS'_0(\mathbf{i}'\ \mathbf{n}'_0) = dS_0(\mathbf{i}\ \mathbf{n}_0), \tag{10}$$

$$dS'(\mathbf{i}'\ \mathbf{n}') = dS(\mathbf{i}\ \mathbf{n}) \tag{11}$$

Inserting from (8)–(11) in (7), we have that the energy E received by A during the interval $(t, t+dt)$ of his (A's) experience, from B, is given by

$$E = (I'_{\nu'}\,d\nu')\,dS_0\,dS\,(\mathbf{i}\ \mathbf{n}_0)(\mathbf{i}\ \mathbf{n})\,dt\frac{(1-V/c)^2}{V^2t^2}, \tag{12}$$

and this energy is received as of frequency ν given by (5) Here $I'_{\nu'}\,d\nu'$ is the 'proper' intensity of radiation of the nebula reckoned in its own frame, by an observer at itself

The important term in (12) is the factor $(1-V/c)^2$, which cancels the similar factor in the denominator of the expression for the particle-density for the simple kinematic model, thereby ensuring no singularity at $V = c$ in the integral for the total background-brightness

NOTE 8 *Impossibility of a discrete distribution of particles in motion satisfying the cosmological principle* (§ 397)

Consider a set of *discrete* particle-observers, in three-dimensional space, finitely separated We have shown in Chapter II that in one-dimensional space it is possible to construct such a system of particle observers satisfying Einstein's cosmological principle, i e such that each describes the system, from his own point of view, in the same way We are now going to prove that it is impossible to construct such a system in three dimensions †

It is sufficient to consider the velocity distribution alone, and it is sufficient to consider the case where all the particles have uniform velocities in the experience of any one of them If possible let such a system exist, and let A, B, C be any three particle observers of the system We propose to establish a contradiction

† In preparing this note I benefited greatly from a conversation with Mr J Hodgkinson, Jesus College, Oxford, for whose interest in the matter I am very grateful

A observes B and C to be moving with certain velocities in his experience Hence C sees B to be moving, in C's experience, with a velocity that can be calculated by transforming B's velocity from A's description of it to C's Since the descriptions of the whole system from A, B, and C are to be identical, A must see a particle moving with the same velocity in A's experience as C sees B to be moving with in C's experience Call this B''_C, the double prime indicating that C's description is involved Similarly, since C sees A to be moving with some velocity in C's experience, B must see some particle moving with the same velocity in his experience Call this particle A'_C, the single prime indicating that a description by B is involved This particle must be seen somewhere by A

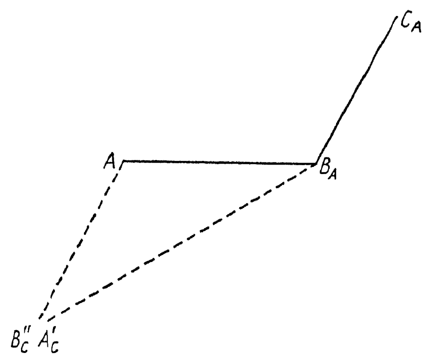


FIG 21 Positions of equivalent particles in a discrete distribution in three dimensions

Thus A should see—

(1) a particle B''_C moving in his (A's) experience with the velocity with which C sees B moving,

(2) a particle A'_C moving with the velocity in B's experience with which C sees A moving

In Newtonian time and space, B''_C and A'_C coincide in velocity, as is readily established When we use the appropriate relations between the times and spaces of the different observers given by the theory of relativity, the velocity of A'_C as judged by A should agree with the velocity of B''_C as judged by A, if a discrete system in three dimensions satisfying the cosmological principle is possible

Take axes of reference such that in A's experience—

A has the velocity $(0, 0, 0)$ or say $\mathbf{w}$,

B has the velocity $(u_1, 0, 0)$ or say $\mathbf{u}$,

C has the velocity $(v_1, v_2, 0)$ or say $\mathbf{v}$

The law of transformation of velocities from A's experience to the experience of an arbitrary particle-observer O' moving with the arbitrary vector-velocity $\mathbf{V}$ is given by the following formula Let $\mathbf{v}$ be the velocity of a particle P in A's experience, $\mathbf{v}'$ the velocity of the same particle in O''s experience Then

$$\mathbf{v}' = \frac{\left[\mathbf{v}-\dfrac{(\mathbf{v}\,\mathbf{V})\mathbf{V}}{\mathbf{V}^2}\right](1-\mathbf{V}^2/c^2)^{\frac{1}{2}}+\left(\dfrac{\mathbf{v}\,\mathbf{V}}{\mathbf{V}^2}-1\right)\mathbf{V}}{1-(\mathbf{v}\,\mathbf{V})/c^2}$$

(This is readily verified to reduce to Einstein's addition-formulae for velocities when $\mathbf{V}$ is taken to be $(V, 0, 0)$)

Hence C describes B as possessing in C's experience the velocity $\mathbf{u}''$ or (u''_1, u''_2, u''_3), where

$$\mathbf{u}'' = \frac{\left[\mathbf{u}-\dfrac{(\mathbf{u}\,\mathbf{v})\mathbf{v}}{v^2}\right](1-v^2/c^2)^{\frac{1}{2}}+\left(\dfrac{\mathbf{u}\,\mathbf{v}}{v^2}-1\right)\mathbf{v}}{1-(\mathbf{u}\,\mathbf{v})/c^2}$$

(Verification $\mathbf{u}'' \sim \mathbf{u}-\mathbf{v}$ for $c \sim \infty$)

Further, since A describes C as moving with the velocity $\mathbf{v}$ in A's experience, C describes A as moving with the velocity $\mathbf{w}'' = -\mathbf{v}$ in C's experience. Hence B sees a particle moving with the velocity $-\mathbf{v}$ in B's experience. This velocity is described by A as $-\mathbf{v}_A$, where (since A has velocity $-\mathbf{u}$ in B's experience)

$$-\mathbf{v}_A = \frac{\left[-\mathbf{v}-\left(\frac{-\mathbf{v}\cdot-\mathbf{u}}{u^2}\right)(-\mathbf{u})\right](1-u^2/c^2)^{\frac{1}{2}}+\left(\frac{(-\mathbf{u})\cdot(-\mathbf{v})}{u^2}-1\right)(-\mathbf{u})}{1-\{(-\mathbf{u})\cdot(-\mathbf{v})\}/c^2}.$$

(Verification. $-\mathbf{v}_A \sim -\mathbf{v}+\mathbf{u}$ for $c \sim \infty$.) If the cosmological principle is satisfied, $\mathbf{u}''$ must be equal to $-\mathbf{v}_A$. This requires

$$\mathbf{u}(1-v^2/c^2)^{\frac{1}{2}}-\mathbf{v}+\frac{\mathbf{u}\cdot\mathbf{v}}{v^2}[1-(1-v^2/c^2)^{\frac{1}{2}}]\mathbf{v}$$
$$= \mathbf{u}-\mathbf{v}(1-u^2/c^2)^{\frac{1}{2}}-\frac{\mathbf{u}\cdot\mathbf{v}}{u^2}[1-(1-u^2/c^2)^{\frac{1}{2}}]\mathbf{u}. \tag{1}$$

The vector condition (1) is equivalent to three scalar conditions.

Multiply (1) vectorially by $\mathbf{v}$. Then

$$(\mathbf{u}\wedge\mathbf{v})(1-v^2/c^2)^{\frac{1}{2}} = (\mathbf{u}\wedge\mathbf{v})-\frac{\mathbf{u}\cdot\mathbf{v}}{u^2}[1-(1-u^2/c^2)^{\frac{1}{2}}](\mathbf{u}\wedge\mathbf{v}).$$

Hence either

$$\mathbf{u}\wedge\mathbf{v} = 0 \tag{2}$$

or

$$(1-v^2/c^2)^{\frac{1}{2}} = 1-\frac{\mathbf{u}\cdot\mathbf{v}}{u^2}[1-(1-u^2/c^2)^{\frac{1}{2}}]. \tag{3}$$

Multiply (1) vectorially by $\mathbf{u}$. Then, by a similar argument, either (2) holds or

$$(1-u^2/c^2)^{\frac{1}{2}} = 1-\frac{\mathbf{u}\cdot\mathbf{v}}{v^2}[1-(1-v^2/c^2)^{\frac{1}{2}}]. \tag{4}$$

Equations (3) and (4) may be regarded as a pair of simultaneous linear equations in $(1-u^2/c^2)^{\frac{1}{2}}$ and $(1-v^2/c^2)^{\frac{1}{2}}$. Solving them we find at once

$$(1-v^2/c^2)^{\frac{1}{2}} = (1-u^2/c^2)^{\frac{1}{2}}, \tag{5}$$

provided
$$1-\left(\frac{\mathbf{u}\cdot\mathbf{v}}{u^2}\right)\left(\frac{\mathbf{u}\cdot\mathbf{v}}{v^2}\right) \neq 0,$$

i.e. provided
$$\mathbf{u}^2\mathbf{v}^2-(\mathbf{u}\cdot\mathbf{v})^2 \neq 0,$$

i.e. provided
$$(\mathbf{u}\wedge\mathbf{v})^2 \neq 0,$$

which is violated only if (2) holds. Hence if (2) does not hold, (5) gives

$$|\mathbf{u}| = |\mathbf{v}|.$$

Call the common value of this modulus V. Then (1) gives

$$\mathbf{u}(1-V^2/c^2)^{\frac{1}{2}}-\mathbf{v}+\frac{\mathbf{u}\cdot\mathbf{v}}{V^2}[1-(1-V^2/c^2)^{\frac{1}{2}}]\mathbf{v}$$
$$= \mathbf{u}-\mathbf{v}(1-V^2/c^2)^{\frac{1}{2}}-\frac{\mathbf{u}\cdot\mathbf{v}}{V^2}[1-(1-V^2/c^2)^{\frac{1}{2}}]\mathbf{u}. \tag{6}$$

Multiply (6) scalarly by $\mathbf{v}$, and put $\mathbf{u}\cdot\mathbf{v} = x$. Then

$$x(1-V^2/c^2)^{\frac{1}{2}}-V^2+x[1-(1-V^2/c^2)^{\frac{1}{2}}]$$
$$= x-V^2(1-V^2/c^2)^{\frac{1}{2}}-\frac{x^2}{V^2}[1-(1-V^2/c^2)^{\frac{1}{2}}],$$

whence
$$0 = \left[V^2-\frac{x^2}{V^2}\right][1-(1-V^2/c^2)^{\frac{1}{2}}].$$

Hence $$x^2 = V^4$$

or $$V^4 = (\mathbf{u}\ \mathbf{v})^2,$$

in which $|\mathbf{u}| = |\mathbf{v}| = V$ This requires either $\mathbf{u} = \mathbf{v}$ or $\mathbf{u} = -\mathbf{v}$ Hence $\mathbf{u} \wedge \mathbf{v} = 0$, or $\mathbf{u}$ is parallel to $\mathbf{v}$ Hence in A's experience B is moving parallel to C For the cosmological principle to be satisfied, B and C must accordingly be collinear with A, and we are driven back to the one dimensional universe of Chapter II

It follows that a discrete system of uniformly moving particle-observers in three dimensions satisfying the cosmological principle cannot exist, Small discrepancies must exist between the different descriptions of the system as seen from the different members of it, of the order of the ratio of the square of the occurring velocities to the square of the velocity of light. We have seen (§§ 500, 501) the overwhelming importance of *uniformity* of velocity of fundamental particles, for this ultimately ensures neither creation nor annihilation of matter, within the experience of an observer, near the expanding frontier. It is not therefore necessary to attempt to generalize the calculation for relatively accelerated fundamental particles, and in any case it is highly improbable that a discrete distribution satisfying the cosmological principle would exist in this case when it does not exist in the case of uniform relative velocity

It now follows that departures from exact identity of description of the system from the different members must be accompanied by what would be called, in a dynamical mode of description, residual fields These will be resolvable into residual relative accelerations, tidal couples, and other couples, which in turn will 'cause' departures from the strictly spherical symmetry of the sub systems associated with the fundamental particles These will be apparent as relative 'proper motions' of their nuclei, the coming into existence of a preferential axis defining an axis of rotation and a plane of equatorial flattening, and other tidal effects. It will remain for the future to investigate what is the minimal extent of departures from satisfaction of the cosmological principle by a discrete system, and to see whether the unavoidable departures are accompanied by rotation, the formation of spiral arms in the manner originally suggested by Jeans, and other regularities It may be observed that the whole 'rest of the universe' is involved Thus we may well provisionally attribute spiral arms to the tidal effect on a rotating system of the whole of the 'rest of the universe', in accordance with Jeans's views Our statistical analysis, as explained in the text, deals only with averages, from which all local distortions and deviations cancel out, hence the average spherical symmetry of a sub-system in the general theory

NAME INDEX

(The references are to pages)

SUBJECT INDEX

(*The references are to numbered sections*)